

OPTIMAL EXACT SIMULATION OF MAX-STABLE AND RELATED RANDOM FIELDS

ZHIPENG LIU, JOSE BLANCHET, A.B. DIEKER, AND THOMAS MIKOSCH

ABSTRACT. We consider the random field

$$M(t) = \sup_{n \geq 1} \{ -\log A_n + X_n(t) \}, \quad t \in T,$$

for a compact set $T \subset \mathbb{R}^m$, where (X_n) is an iid sequence of centered Gaussian random fields on T and $0 < A_1 < A_2 < \dots$ are the arrivals of a general renewal process on $(0, \infty)$, independent of (X_n) . In particular, a large class of max-stable random fields with Gumbel marginals have such a representation. Assume that one needs $c(d) = c(\{t_1, \dots, t_d\})$ function evaluations to sample X_n at d locations $t_1, \dots, t_d \in T$. We provide an algorithm which, for any $\epsilon > 0$, samples $M(t_1), \dots, M(t_d)$ with complexity $o(c(d) d^\epsilon)$. Moreover, if X_n has an a.s. converging series representation, then M can be a.s. approximated with error δ uniformly over T and with complexity $O(1/(\delta \log(1/\delta))^{1/\alpha})$, where α relates to the Hölder continuity exponent of the process X_n (so, if X_n is Brownian motion, $\alpha = 1/2$).

September 21, 2016

1. INTRODUCTION

Let X be a centered Gaussian random field on a set $T \subseteq \mathbb{R}^m$, $m \geq 1$ and consider a sequence (X_n) of independent and identically distributed copies of X . In addition, let (A_n) be a renewal sequence independent of (X_n) . Under mild regularity conditions on the X , we will provide an efficient Monte-Carlo algorithm for sampling the field

$$(1.1) \quad M(t) = \sup_{n \geq 1} \{ -\log A_n + X_n(t) + \mu(t) \}, \quad t \in T,$$

where $\mu : T \rightarrow \mathbb{R}$ is a bounded function.

We will design and analyze an algorithm for the exact simulation of

$$M(t_1), \dots, M(t_d) \quad \text{for any choice of distinct locations } t_1, \dots, t_d \in T,$$

and we will show that, in some sense, this algorithm is asymptotically optimal as $d \rightarrow \infty$.

The algorithm proposed here shaves off a factor of order (nearly) d from the running time of any of the existing exact sampling procedures. In particular, we will show, under mild boundedness assumptions on X , that it is as hard to sample $(M(t_i))_{i=1, \dots, d}$ as it is to sample $(X(t_i))_{i=1, \dots, d}$. Therefore, at least from a simulation point of view, it is not more difficult to work with M than with X . More precisely, if it takes $O(c(d))$ units of computing time to sample X at d distinct locations $t_1, \dots, t_d \in T$, then, for any given $\epsilon > 0$, it takes $o(c(d) d^\epsilon)$ units to sample M at the same locations; see Theorem 2.1 for a precise formulation.

We illustrate this result by considering fractional Brownian motion X on $T = [0, 1]$. Using the circulant-embedding method (see [3], Section XI.3), we have $c(d) = O(d \log d)$ provided we sample at the dyadic points $t_i = i/2^{-m}$ for $i = 1, \dots, 2^m = d$ (which we call dyadic points at level d).

1991 *Mathematics Subject Classification*. Primary; Secondary.

Key words and phrases. Keywords:

Support from NSF grant DMS-132055 is gratefully acknowledged by J. Blanchet. Thomas Mikosch's research is partly supported by the Danish Research Council Grant DFF-4002-00435 "Large random matrices with heavy tails and dependence". A.B. Dieker gratefully acknowledges support from NSF grant CMMI-1252878.

In the case of Brownian motion, one even has $c(d) = O(d)$, corresponding to the simulation of d independent Gaussian random variables. Thus, in the case of fractional Brownian motion on $[0, 1]$ we provide an algorithm for sampling M at the dyadic points at level d in $[0, 1]$ with complexity $o(d^{1+\epsilon})$ for any $\epsilon > 0$; see [3, Sec. XI.6].

Moreover, if X has a series representation a.s. converging uniformly on T (such as the Lévy-Ciesielski representation for Brownian motion, see [27, Sec. 3.1], we also propose an approximate simulation procedure for M with a user-defined (deterministic) bound on the error which holds with probability one uniformly throughout T . More precisely, for any $\delta > 0$, the procedure that we present outputs an approximation M_δ to M such that

$$(1.2) \quad \sup_{t \in T} |M(t) - M_\delta(t)| \leq \delta \quad \text{a.s.}$$

The results concerning (1.2) are reported in Theorem 7.4. The method of designing a family $(M_\delta)_{\delta>0}$ such that (1.2) holds is known as *Tolerance Enforced Simulation* (TES) or δ -strong simulation; see [9] and [26] for details. Note that a TES algorithm enforces a strong (almost sure) guarantee without knowledge of any specific set of sampling locations. This is a feature which distinguishes TES from more traditional algorithms in the broad literature on simulation of random fields and processes.

As will be explained later, the evaluation of $M_\delta(t)$ for fixed t takes $O(1)$ units of computing time while the construction of the process M_δ will often take $O(1/(\delta \log(1/\delta))^2)$ units. The latter result holds under assumptions on the convergence of the series representation of X which, in particular, are satisfied for Brownian motion X . In the latter case, the proposed procedure achieves a complexity of order $O(d)$ for the exact sampling of M on the dyadic points at level d (because the series truncated at level d is exact on the dyadic points at level d). Therefore, the exact sampling procedure based on Theorem 7.4 applied to the dyadic points at level d is optimal because it takes $O(d)$ computational cost to sample X at d dyadic points. Moreover, the convergence rate in (1.2) is also optimal in the Brownian case. In order to obtain a uniform error of order $O(\delta)$, one requires to discretize Brownian motion using a grid of size $O(1/(\delta \log(1/\delta))^2)$; see [5].

Our results are mainly motivated by application to the simulation of max-stable random fields. Indeed, if (A_i) is the arrival sequence of a unit rate Poisson process on $(0, \infty)$, M is a *max-stable process* in the sense of de Haan [21]. This means, in particular, that the distribution of $M(t)$ for any fixed $t \in T$ has a Gumbel distribution which is one of the max-stable distributions. The latter class of distributions consists of the non-degenerate limit distributions for the suitably centered and scaled partial maxima of an iid sequence; see for example [19]. The non-Gumbel max-stable processes with Fréchet or Weibull marginals are obtained from the representation (1.1) by suitable monotone transformations. We also mention that de Haan [21] already proved that max-stable processes with Gumbel marginals have representation (1.1), where X may have a rather general dependence structure not restricted to Gaussian X . However, the case of Gaussian X has attracted major attention. The case of Brownian X was treated in [15] and it is known as the *Brown-Resnick process*. Other examples include the case of a Gaussian process X defined on a multidimensional set T often referred to as *Smith model*, which is used in environmental applications for modeling storm profiles; for example, see [29]. General characterizations, including spectral representations and further properties, have been obtained as well; for example, see [28] and [23]. However, the explicit joint distribution of the max-stable process is in general not tractable. Because max-stable processes are generated as weak limits of maxima of iid random fields, max-stable models are particularly suited for modeling extremal events in spatiotemporal contexts. These include a wide range of applications of environmental type, for example, extreme rainfall [20] and extreme temperature [31].

Recently, several exact sampling procedures for M have been proposed and studied in the literature. In [17], an elegant and easy-to-implement procedure was proposed for the case in which

X has stationary increments. Such a procedure has a computational complexity at least of order $O(c(d)d)$; see Proposition 4 in [18]. So, for example, if X is fractional Brownian motion, the procedure takes at least $O(d^2 \log d)$ units of computing time to produce d dyadic points of M in $[0, 1]$.

Another exact simulation method for M was recently proposed in [18]. It also has complexity $O(c(d)d)$ (see Proposition 4 in [18]), thus the procedure in [18] takes $O(d^2)$ for fractional Brownian motion (neglecting the contribution of logarithmic factors). This method is based on the idea of simulating the extremal functions. It is completely different from the approach taken here. Additional work concentrates on max-stable processes which satisfy special characteristics. For example, [28] proposed an exact simulation algorithm for the Schlather model under suitable uniformity conditions, and [25] discussed the exact simulation strategy for the max-stable process under certain uniform bounds on the underlying spectral representation of the model.

An important difference between our method and those in [17] and [18] is the following: Both [17] and [18] take advantage of representations or structures which allow to truncate the infinite max-convolution in (1.1) while preserving the simple Gaussian structure of the number of terms in the truncation. Because the simple structure of these terms is preserved, the number of terms in the truncation increases at least linearly in d . In contrast, we are able to truncate the number of terms in the infinite max-convolution uniformly in d . While the terms in the truncation have a slightly more complex structure (they are no longer iid Gaussian), they are still quite tractable from a simulation standpoint.

This paper is organized as follows: In Section 2 we present our main result and in Section 3 we discuss our general strategy, based on *milestone events* or *record-breakers*. The record-breaking strategy is illustrated in Section 4 in the setting of random walks, which is needed in our context due to the presence of (A_n) in M . Then we apply the record-breaking strategy to the setting of maxima of Gaussian random vectors with focus on Section 5: This section describes the main algorithmic developments of the paper. A complexity analysis is performed in Section 6. We introduce and analyze a TES algorithm in Section 7. Finally, in Section 8, we conclude our paper with a series of empirical comparison results.

2. MAIN RESULT

This section provides a formal statement of the main result and its underlying assumptions. We assume that $(A_n)_{n \geq 0}$ is a renewal sequence, as mentioned in the Introduction. In particular, $A_0 = 0$, and $A_n = \tau_1 + \dots + \tau_n$, $n \geq 1$, where (τ_i) is an iid sequence of positive random variables, independent of (X_n) . In what follows, we τ a for generic element of the sequence (τ_i) .

We introduce the following technical assumptions applicable to (A_n) :

- A1) For any $\gamma < \mathbb{E}\tau_1$, there exists some $\theta_\gamma > 0$ such that $\mathbb{E}[\exp(\theta_\gamma(\gamma - \tau_1))] = 1$.
- A2) It is possible to sample step sizes under the nominal probability measure as well as under the exponentially tilted distribution

$$\mathbb{E}[\exp(\theta_\gamma(\gamma - \tau_1))\mathbf{1}(\tau_1 \in dt)].$$

We also introduce the following assumptions applicable to the Gaussian field $(X(t))_{t \in T}$ and the function μ .

- B1) We assume that $\mathbb{E}[X(t)] = 0$.
- B2) Assume $|\mu(t)| < \infty$ for any $t \in T$.
- B3) Suppose that $\mathbb{E}[\exp(p \sup_{t \in T} X(t))] < \infty$ for any $p \geq 1$.

By Borell's inequality [1, Thm. 2.1.1], a sufficient condition for B3) is

$$\text{Var}(X(s) - X(t)) \leq c|s - t|^\beta$$

for any $s, t \in T$ and some $c > 0$, $\beta > 0$. We define $\sigma^2(t) = \text{Var}(X(t))$, it follows from Assumption B3) that

$$\sup_{t \in T} \sigma^2(t) = \text{Var}(X(t)) = \sup_{t \in T} \mathbb{E}[X(t)^2] \leq \mathbb{E}[\sup_{t \in T} X(t)^2] < \infty.$$

We also assume that sampling $(X(t_i))_{i=1, \dots, d}$ costs $c(\{t_1, \dots, t_d\}) \geq d$ units of operations. In this paper, a single operation can be any single arithmetic operation, generating a uniform random variable, calculating a Gaussian cumulative probability function, comparing any two numbers, or retrieving a Gaussian quantile value. For simplicity in the notation, we shall simply write $c(d) = c(\{t_1, \dots, t_d\})$. The locations t_1, \dots, t_d will be assumed given throughout our development.

The following is our performance guarantee for our final algorithm, **Algorithm M**, presented in Section 6. A crucial part of the theorem is that the points t_1, \dots, t_d for any $d \geq 1$ lie in a fixed set T .

Theorem 2.1. *Suppose that Assumptions A1), A2), and B1)–B3) hold. Let R be the total number of operations in the execution of **Algorithm M**. Then, **Algorithm M** outputs $M(t_1), \dots, M(t_d)$ without any bias, and $\mathbb{E}[R^p] = o(d^\epsilon c(d)^p)$ for any $p \geq 1$ and $\epsilon > 0$.*

3. BUILDING BLOCKS FOR OUR ALGORITHM

This section serves as a roadmap for the algorithmic elements behind our approach. We start with a few definitions:

$$\overline{X}_n = \max_{i=1, \dots, d} X_n(t_i), \quad \underline{X}_n = \min_{i=1, \dots, d} X_n(t_i).$$

We shall use \overline{X} and \underline{X} to denote generic copies of \overline{X}_n and \underline{X}_n , respectively. We also set $\overline{\mu} = \max_{i=1, \dots, n} \mu(t_i)$ and $\underline{\mu} = \min_{i=1, \dots, n} \mu(t_i)$.

Our algorithm relies on three random times to be defined next. The first is $N_X = N_X(a, C) < \infty$, defined for some $a \in (0, 1)$ and $C > 0$ to be chosen in the sequel, which satisfies, for all $n > N_X$,

$$\overline{X}_n \leq a \log n + C.$$

The second is $N_A = N_A(\gamma) < \infty$ for some $\gamma < \mathbb{E}A_1$, which satisfies, for all $n > N_A$,

$$(3.1) \quad A_n \geq \gamma n.$$

The third is $N_a = N_a(\gamma, a, C)$, which satisfies, for all $n > N_a$,

$$(3.2) \quad n\gamma \geq A_1 n^a \exp(C - \underline{X}_1).$$

By successively applying the preceding three displays, we find that for $n > N := \max(N_A, N_X, N_a)$ and any $t = t_1, \dots, t_d$ we have

$$\begin{aligned} -\log A_n + X_n(t) &\leq -\log A_n + \overline{X}_n \\ &\leq -\log A_n + a \log n + C \\ &\leq -\log(n\gamma) + a \log n + C \\ &\leq -\log A_1 + \underline{X}_1 \\ &\leq -\log A_1 + X_1(t). \end{aligned}$$

We conclude that, for $t \in \{t_1, \dots, t_d\}$,

$$(3.3) \quad \sup_{n \geq 1} \{-\log A_n + X_n(t) + \mu(t)\} = \max_{1 \leq n \leq N} \{-\log A_n + X_n(t) + \mu(t)\},$$

and thus we can sample $M(t_1), \dots, M(t_d)$ with computational complexity $Nc(\{t_1, \dots, t_d\})$ plus the overhead to identify N_A and N_X .

From an algorithmic point of view, the key is the simulation of the random variables N_X , N_A , and N_a . For instance, (3.3) shows that we need to sample N_A jointly with the sequence $(A_n)_{n \leq N_A}$ and in fact even jointly with the longer sequence $(A_n)_{n \leq N}$ as well as $(X_n)_{n \leq N}$.

Remark 3.1. *In fact, we can remove assumptions A1) and A2) without loss of generality. First, observe that for any $r > 0$ we have that $\tau_i(r) = \min(\tau_i, r) \leq \tau_i$ and, therefore,*

$$A_n(r) = \tau_1(r) + \dots + \tau_n(r) \leq A_n.$$

Moreover, we can select $r > 0$ so that $\gamma < \mathbb{E}[\tau_i(r)] < \mathbb{E}[\tau_i]$, which means that we can use $(A_n(r))_{n \geq 1}$ in order to find N_A satisfying

$$A_n > A_n(r) > n\gamma,$$

which implies (3.1). Because $0 \leq \tau_n(r) \leq r$, the moment generating function of $\tau_n(r)$ exists on the whole real line. By convexity, one can always choose θ_γ which satisfies $\mathbb{E}[\exp(\theta_\gamma(\gamma - \tau_1(r)))] = 1$, as long as $\text{Var}(\tau_i(r)) > 0$ (i.e. if $\tau_i > 0$ is non-deterministic, by choosing $r > 0$ large enough). If τ_i is deterministic, the strategy can be implemented directly, that is, we can simply select N_A deterministic.

Given our previous discussion, we might concentrate on how to sample from an exponentially tilted distribution of a random variable with compact support, which may require evaluating the moment generating function in closed form. Sampling from an exponentially tilted distribution is straightforward for random variables with finite support. So, the strategy can be implemented for $[\tau_i(r)\Delta]/\Delta < \tau_i(r)$, picking $\Delta > 0$ sufficiently small so that $\mathbb{E}[[\tau_i(r)\Delta]/\Delta] > \gamma$. Once $[\tau_i(r)\Delta]$ is sampled we can easily simulate $\tau_i(r)$ using acceptance/rejection. The details of this idea are explained in [13].

4. SAMPLING A RANDOM WALK UP TO A LAST PASSAGE TIME

Given a random walk $(S_n)_{n \geq 0}$ starting from the origin with negative drift, this section reviews an algorithm from [14] for finding a random time N_S such that $S_n < 0$ for all $n > N_S$. Our aim is to develop a sampling algorithm for $(S_1, \dots, S_{N_S+\ell})$ for any fixed $\ell \geq 0$. Our discussion here provides a simpler version of the algorithm in [14] and allows us to provide a self-contained development of the whole procedure to sample $M(t_1), \dots, M(t_d)$.

The algorithm is based on alternately sampling upcrossings and downcrossings of the level 0. We write $\xi_0^+ = 0$ and, for $i \geq 1$, we recursively define

$$\xi_i^- = \begin{cases} \inf\{n \geq \xi_{i-1}^+ : S_n < 0\} & \text{if } \xi_{i-1}^+ < \infty \\ \infty & \text{otherwise} \end{cases}$$

together with

$$\xi_i^+ = \begin{cases} \inf\{n \geq \xi_i^- : S_n \geq 0\} & \text{if } \xi_i^- < \infty \\ \infty & \text{otherwise.} \end{cases}$$

As usual, in these definitions the infimum of an empty set should be interpreted as ∞ . Writing

$$N_S = \sup\{\xi_n^- : \xi_n^- < \infty\},$$

we have by construction $S_n \leq 0$ for $n > N_S$ (it even holds for $n = N_S$ but we do not include this for expository reasons). The random variable $N_S - 1$ is an upward last passage time:

$$N_S - 1 = \sup\{n \geq 0 : S_n \geq 0\}.$$

Note that $0 \leq N_S < \infty$ almost surely under \mathbb{P} because $(S_n)_{n \geq 0}$ starts at the origin and has negative drift.

In order to introduce our algorithm, we need more notation. We write \mathbb{P}_x for the distribution of the random walk starting from $x \in \mathbb{R}$, so that $\mathbb{P} = \mathbb{P}_0$. We assume the existence of *Cramér's root*, $\theta > 0$, satisfying $\mathbb{E}(\exp(\theta S_1)) = 1$. Also assume that we can sample a random walk starting from x

under \mathbb{P}_x^θ , which is defined with respect to \mathbb{P}_x through an exponential change of measure: On the σ -field generated by S_1, \dots, S_n we have

$$\frac{d\mathbb{P}_x^\theta}{d\mathbb{P}_x} = \exp(-\theta(S_n - x)).$$

Under \mathbb{P}_x^θ , the random walk (S_n) has positive drift.

The rest of this section is organized as follows: Section 4.1 first discusses sampling of downcrossing and upcrossing segments of the random walk. Section 4.2 discusses how to sample beyond N_S . Section 4.3 presents our full algorithm for sampling $(S_1, \dots, S_{N_S+\ell})$.

4.1. Downcrossings and upcrossings. To introduce the algorithm, we first need the following definitions:

$$\tau^- = \inf\{n \geq 0 : S_n < 0\}, \quad \tau^+ = \inf\{n \geq 0 : S_n \geq 0\}.$$

For $x \geq 0$, it is immediate that we can sample a downcrossing segment S_1, \dots, S_{τ^-} under \mathbb{P}_x due to the negative drift, and we record this for later use in a pseudocode function. *Throughout this paper, ‘sample’ in pseudocode stands for ‘sample independently of anything that has been sampled already.’*

Function SAMPLEDOWNCROSSING(x): **Samples** (S_1, \dots, S_{τ^-}) **under** \mathbb{P}_x **for** $x \geq 0$

Step 1: Return sample (S_1, \dots, S_{τ^-}) under \mathbb{P}_x .

Step 2: EndFunction

Sampling an upcrossing segment is much more challenging because it is possible that $\tau^+ = \infty$, so an algorithm needs to be able to detect this event within a finite amount of computing resources. For this reason, we understand sampling an upcrossing segment under \mathbb{P}_x for $x < 0$ to mean that an algorithm outputs (S_1, \dots, S_{τ^+}) if $\tau^+ < \infty$, and otherwise it outputs ‘degenerate.’

Our algorithm is based on importance sampling and exponential tilting, techniques that are widely used for rare event simulation [3, p. 164]. Under Assumption A2), it is well-known that $\mathbb{E}_x^\theta[\tau^+] < \infty$; for instance, see [2, Cor. 4.4]. In particular, the expected time to simulate (S_1, \dots, S_{τ^+}) is finite under \mathbb{P}_x^θ for any $x < 0$.

The following proposition is the key to our algorithm.

Proposition 4.1. *Let $x < 0$. Suppose there exists some $\theta > 0$ with $\mathbb{E}[\exp(\theta S_1)] = 1$. With U being a standard uniform random variable independent of (S_n) under \mathbb{P}_x^θ , we have the following:*

- (1) *The law of $\mathbf{1}(\tau^+ < \infty)$ under \mathbb{P}_x equals the law of $\mathbf{1}(U \leq \exp(-\theta(S_{\tau^+} - x)))$ under \mathbb{P}_x^θ .*
- (2) *The law of τ^+ given $\tau^+ < \infty$ under \mathbb{P}_x equals the law of τ^+ given $U \leq \exp(-\theta(S_{\tau^+} - x))$ under \mathbb{P}_x^θ .*
- (3) *For any $k \geq 1$, the law of (S_1, \dots, S_k) given $\tau^+ = k$ under \mathbb{P}_x equals the law of (S_1, \dots, S_k) given $U \leq \exp(-\theta(S_{\tau^+} - x))$ and $\tau^+ = k$ under \mathbb{P}_x^θ .*

Proof. For any integer $k \geq 1$ and Borel sets B_1, B_2, \dots, B_k , we have

$$\begin{aligned} \mathbb{P}_x(S_1 \in B_1, \dots, S_k \in B_k, \tau^+ = k) \\ &= \mathbb{E}_x^\theta[\exp(-\theta(S_k - x)) \mathbf{1}(S_1 \in B_1, \dots, S_k \in B_k, \tau^+ = k)] \\ &= \mathbb{E}_x^\theta[\mathbf{1}(U \leq \exp(-\theta(S_{\tau^+} - x))) \mathbf{1}(S_1 \in B_1, \dots, S_k \in B_k, \tau^+ = k)]. \end{aligned}$$

All claims are elementary consequences of this identity, upon noting that $\tau^+ < \infty$ under \mathbb{P}_x^θ . \square

This proposition immediately yields the following algorithm.

Function SAMPLEUPCROSSING(x): **Samples** (S_1, \dots, S_{τ^+}) **under** \mathbb{P}_x **for** $x < 0$

Step 1: $S \leftarrow$ sample (S_1, \dots, S_{τ^+}) under \mathbb{P}_x^θ

Step 2: $U \leftarrow$ sample a standard uniform random variable

Step 3: If $U \leq \exp(-\theta(S_{\tau^+} - x))$

Step 4: Return S
 Step 5: Else
 Step 6: Return 'degenerate'
 Step 7: EndIf
 Step 8: EndFunction

4.2. **Beyond N_S .** We next describe how to sample (S_1, \dots, S_n) from \mathbb{P}_x conditionally on $\tau^+ = \infty$ for $x < 0$. Because $\tau^+ = \infty$ is equivalent to $\sup_{k \leq \ell} S_k < 0$ and $\sup_{k > \ell} S_k < 0$ for any $\ell \geq 1$, after sampling S_1, \dots, S_ℓ , by the Markov property we can use $\text{SAMPLEUPCROSSING}(S_\ell)$ to verify whether or not $\sup_{k > \ell} S_k < 0$. This observation immediately yields an acceptance/rejection algorithm that achieves our goal.

Function $\text{SAMPLEWITHOUTRECORDS}(x, \ell)$: **Samples** (S_1, \dots, S_ℓ) **from** \mathbb{P}_x **given** $\tau^+ = \infty$ **for** $\ell \geq 1, x < 0$

Step 1: Repeat
 Step 2: $S \leftarrow \text{sample}(S_1, \dots, S_\ell)$ under \mathbb{P}_x
 Step 3: Until $\sup_{1 \leq k \leq \ell} S_k < 0$ and $\text{SAMPLEUPCROSSING}(S_\ell)$ is 'degenerate'
 Step 4: Return S
 Step 5: EndFunction

4.3. **Sampling a random walk until a last passage time.** We summarize our findings in this section in our full algorithm for sampling $(S_0, \dots, S_{N_S + \ell})$ under \mathbb{P} given some $\ell \geq 0$. The validity of the algorithm is a direct consequence of the strong Markov property.

Algorithm S: Samples $(S_0, \dots, S_{N_S + \ell})$ **under** \mathbb{P} **for** $\ell \geq 0$. # We use S_{end} to denote the last element of S .

Step 1: $S \leftarrow [0]$
 Step 2: Repeat
 Step 3: DowncrossingSegment $\leftarrow \text{SAMPLEDOWNCROSSING}(S_{\text{end}})$
 Step 4: $S \leftarrow [S, \text{DowncrossingSegment}]$
 Step 5: UpcrossingSegment $\leftarrow \text{SAMPLEUPCROSSING}(S_{\text{end}})$
 Step 6: If UpcrossingSegment is not 'degenerate'
 Step 7: $S \leftarrow [S, \text{UpcrossingSegment}]$
 Step 8: EndIf
 Step 9: Until UpcrossingSegment is 'degenerate'
 Step 10: If $\ell > 0$
 Step 11: $S \leftarrow [S, \text{SAMPLEWITHOUTRECORDS}(S_{\text{end}}, \ell)]$
 Step 12: EndIf

5. RECORD-BREAKER TECHNIQUE FOR THE MAXIMUM OF A GAUSSIAN FIELD

After the excursion to random walks in Section 4 we return to the main theme of this paper. In particular, we stick to the notation and assumptions of Section 1-3. Define $\eta_0 = 0$. Let $(X_n)_{n \geq 1}$ be iid copies of X and define, for $i \geq 1$, a sequence of *record-breaking times* (η_i) through

$$\eta_i = \begin{cases} \inf\{n > \eta_{i-1} : \bar{X}_n > a \log n + C\} & \text{if } \eta_{i-1} < \infty \\ \infty & \text{otherwise.} \end{cases}$$

It is the aim of this section to develop a sampling algorithm for $(X_1, \dots, X_{N_X + \ell})$ for any fixed $\ell \geq 0$, where

$$N_X = \max\{\eta_i : \eta_i < \infty\}.$$

Here and in what follows, we write X_i for sample path at the given points $t_1, \dots, t_d \in T$. Section 5.1 first discusses an algorithm to sample (X_n) up to a single record. For this algorithm to

work, n_0 needs to be large enough so that $\mathbb{P}(\bar{X} > a \log n + C)$ is controlled for every $n > n_0$; the choice of n_0 is also discussed in Section 5.1. Section 5.2 describes how to sample (X_n) beyond the last record-breaking time. Section 5.3 presents our algorithm for sampling $(X_1, \dots, X_{N_X+\ell})$.

5.1. Breaking a single record. Fix some integer $n_0 \geq 0$, and define

$$T_{n_0} = \inf\{k \geq 1 : \bar{X}_k > a \log(n_0 + k) + C\}.$$

We describe an algorithm that outputs ‘degenerate’ if $T_{n_0} = \infty$ and $(X_1, \dots, X_{T_{n_0}})$ if $T_{n_0} < \infty$. Ultimately, the strategy is based on acceptance / rejection. We will eventually sample T_{n_0} given that $T_{n_0} < \infty$ using a suitable random variable K , with probability mass function as a proposal g_{n_0} . However, in order to apply this acceptance / rejection strategy, we need to introduce auxiliary sampling distributions.

Our algorithm makes use of a measure $\mathbb{P}^{(n)}$ that is designed to appropriately approximate the conditional distribution of X given $\bar{X} > a \log n + C$, which is defined through

$$\frac{d\mathbb{P}^{(n)}}{d\mathbb{P}}(x) = \frac{\sum_{i=1}^d \mathbf{1}(x(t_i) > a \log n + C)}{\sum_{i=1}^d \mathbb{P}(X(t_i) > a \log n + C)}.$$

Because with any given index $j \in \{1, \dots, d\}$, the random vector $X - w^j X(t_j)$ is independent of $X(t_j)$, where w^j is defined as $w^j(t) = \text{Cov}(X(t), X(t_j)) / \text{Var}(X(t_j))$ for $t \in \{t_1, \dots, t_d\}$, one readily verifies that the following algorithm outputs samples from $\mathbb{P}^{(n)}$. Note that $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

Function CONDITIONEDSAMPLEX (a, C, n): **Samples** X **from** $\mathbb{P}^{(n)}$

Step 1: $\nu \leftarrow$ sample with probability mass function

$$\mathbb{P}(\nu = j) = \frac{\mathbb{P}(X(t_j) > a \log n + C)}{\sum_{i=1}^d \mathbb{P}(X(t_i) > a \log n + C)}$$

Step 2: $U \leftarrow$ sample a standard uniform random variable

Step 3: $X(t_\nu) \leftarrow \sigma(t_\nu) \Phi^{-1} \left(U + (1 - U) \Phi \left(\frac{a \log n + C}{\sigma(t_\nu)} \right) \right)$ # Conditions on $X(t_\nu) > a \log n + C$

Step 4: $Y \leftarrow$ sample of X under P

Step 5: Return $Y - w^\nu Y(t_\nu) + X(t_\nu)$

Step 6: EndFunction

We are now ready to see how CONDITIONEDSAMPLEX is used to sample until the first record.

Function SAMPLESINGLERECORD (a, C, n_0, n_1): **Samples** $(X_1, \dots, X_{T_{n_1}})$ **for** $a \in (0, 1), C \in \mathbb{R}, n_1 \geq n_0 \geq 0$

Step 1: $K \leftarrow$ sample from pmf g_{n_0}

Step 2: $(X_1, \dots, X_{K-1}) \leftarrow$ iid sample under \mathbb{P}

Step 3: $X_K \leftarrow$ CONDITIONEDSAMPLEX($a, C, n_1 + K$)

Step 4: $U \leftarrow$ sample a standard uniform random variable

Step 5: If $\bar{X}_k \leq a \log(n_1 + k) + C$ for $k = 1, \dots, K - 1$ and $U g_{n_0}(K) \leq d\mathbb{P}/d\mathbb{P}^{(n_1+K)}(X_K)$

Step 6: Return (X_1, \dots, X_K)

Step 7: Else

Step 8: Return ‘degenerate’

Step 9: EndIf

Step 10: EndFunction

The following proposition shows that SAMPLESINGLERECORD achieves the desired goal.

Proposition 5.1. *Suppose $\sum_{i=1}^d \mathbb{P}(X(t_i) > a \log(n_0 + k) + C) \leq g_{n_0}(k)$ for $k \geq 1$. For $n_1 \geq n_0$, if $(\tilde{X}_1, \dots, \tilde{X}_{\tilde{T}})$ has the distribution of the output of SAMPLESINGLERECORD conditioned on not being ‘degenerate,’ then we have*

- (1) *the algorithm SAMPLESINGLERECORD returns ‘degenerate’ with probability $\mathbb{P}(T_{n_1} = \infty)$,*
- (2) *the length \tilde{T} has the same distribution as T_{n_1} , and*
- (3) *the distribution of $(\tilde{X}_1, \dots, \tilde{X}_{\tilde{T}})$ given $\tilde{T} = \ell$ is the same as the distribution of (X_1, \dots, X_ℓ) given $T_{n_1} = \ell$.*

Proof. Write $A_m = \{x \in \mathbb{R}^d : \max_i x_i > a \log(n_1 + m) + C\}$ for $m \geq 1$. For $x^{(1)} \in A_1^c, \dots, x^{(k-1)} \in A_{k-1}^c$ and $x^{(k)} \in A_k$, we have

$$\begin{aligned}
& \mathbb{P}(\tilde{X}_1 \in dx^{(1)}, \dots, \tilde{X}_{k-1} \in dx^{(k-1)}, \tilde{X}_k \in dx^{(k)}, \tilde{T} = k) \\
&= \mathbb{P}(K = k) \mathbb{P}(X \in dx^{(1)}) \dots \mathbb{P}(X \in dx^{(k-1)}) \\
&\quad \times \mathbb{P}^{(n_1+k)} \left(U g_{n_0}(k) \leq \frac{d\mathbb{P}}{d\mathbb{P}^{(n_1+k)}}(X), X \in dx^{(k)} \right) \\
&= \mathbb{P}(K = k) \mathbb{P}(X \in dx^{(1)}) \dots \mathbb{P}(X \in dx^{(k-1)}) \\
&\quad \times \mathbb{E}^{(n_1+k)} \left(\frac{1}{g_{n_0}(k)} \frac{d\mathbb{P}}{d\mathbb{P}^{(n_1+k)}}(X), X \in dx^{(k)} \right) \\
&= g_{n_0}(k) \mathbb{P}(X \in dx^{(1)}) \dots \mathbb{P}(X \in dx^{(k-1)}) \frac{\mathbb{P}(X \in dx^{(k)})}{g_{n_0}(k)} \\
&= \mathbb{P}(X_1 \in dx^{(1)}, \dots, X_k \in dx^{(k)}, T_{n_1} = k),
\end{aligned}$$

and both claims follow from this identity. The second equality follows from the assumption, which implies that $d\mathbb{P}/d\mathbb{P}^{(n_1+k)}(x)/g_{n_0}(k)$ is bounded by 1 for all $k \geq 1$ and $x \in \mathbb{R}^d$. \square

Choosing n_0 and the density g_{n_0} . We start with g_{n_0} , guided by the condition in Proposition 5.1 and the requirement that we need to sample from g_{n_0} . A natural choice is, for $k \geq 1$,

$$(5.1) \quad g_{n_0}(k) = \frac{\int_{k-1}^k \phi((a \log(n_0 + s) + C)/\bar{\sigma}) ds}{\int_0^\infty \phi((a \log(n_0 + s) + C)/\bar{\sigma}) ds},$$

where $\phi(\cdot)$ is the density function of the standard normal distribution, $\bar{\sigma}^2 = \max_{i=1, \dots, d} \text{Var}(X(t_i))$. The following lemma resolves the sampling question.

Lemma 5.2. *Let U be a uniform random variable on $(0, 1)$. The quantity*

$$\left\lceil \exp \left\{ \frac{\bar{\sigma}^2}{a^2} - \frac{C}{a} + \frac{\bar{\sigma}}{a} \bar{\Phi}^{-1} \left(U \bar{\Phi} \left(\frac{a \log n_0 + C}{\bar{\sigma}} - \frac{\bar{\sigma}}{a} \right) \right) \right\} - n_0 \right\rceil$$

is a random variable with probability mass function g_{n_0} , where $\lceil \cdot \rceil$ is the round-up operator, $\bar{\Phi}(\cdot) = 1 - \Phi(\cdot)$, and $\bar{\Phi}^{-1}$ is the inverse of $\bar{\Phi}$.

Proof. Write $f_{n_0}(U)$ for the expression inside the exponential operator. For $k \geq 1$, we have

$$\begin{aligned}
\mathbb{P}(\lceil \exp(f_{n_0}(U)) - n_0 \rceil \geq k) &= \mathbb{P}(f_{n_0}(U) > \log(n_0 + k - 1)) \\
&= \frac{\bar{\Phi}((a \log(n_0 + k - 1) + C)/\bar{\sigma} - \bar{\sigma}/a)}{\bar{\Phi}((a \log n_0 + C)/\bar{\sigma} - \bar{\sigma}/a)},
\end{aligned}$$

so it remains to show that this equals

$$\sum_{m \geq k} g_{n_0}(m) = \frac{\int_{n_0+k-1}^\infty \phi((a \log x + C)/\bar{\sigma}) dx}{\int_{n_0}^\infty \phi((a \log x + C)/\bar{\sigma}) dx}.$$

To see this, we note that, for $y > 0$,

$$\begin{aligned} \int_y^\infty \phi((a \log(x) + C)/\bar{\sigma}) dx &= \frac{1}{\sqrt{2\pi}} \int_{\log y}^\infty \exp\left(-\frac{(at + C)^2}{2\bar{\sigma}^2} + t\right) dt \\ &= \frac{e^{-C/a}}{\sqrt{2\pi}\phi(\bar{\sigma}/a)/(\bar{\sigma}/a)} \times \bar{\Phi}((a \log y + C)/\bar{\sigma} - \bar{\sigma}/a) \end{aligned}$$

and we thus obtain the claim. \square

The next lemma shows that, for large enough n_0 , this choice of g_{n_0} ensures that the condition from Proposition 5.1 is satisfied. It also shows how $\mathbb{P}(T_{n_1} < \infty)$ for $n_1 \geq n_0$ can be controlled explicitly.

Proposition 5.3. *Let $\delta \in (0, 1)$ be given. If n_0 satisfies $a \log n_0 + C \geq \bar{\sigma}$ and*

$$de^{-C/a} \bar{\Phi}\left(\frac{a \log n_0 + C}{\bar{\sigma}} - \frac{\bar{\sigma}}{a}\right) \leq \delta \sqrt{2\pi} \frac{\phi(\bar{\sigma}/a)}{\bar{\sigma}/a},$$

then the condition in Proposition 5.1 is satisfied and `SAMPLESINGLERECORD`(a, C, n_0, n_1) returns ‘degenerate’ with probability at least $1 - \delta$.

Proof. Because $\bar{\Phi}(x) \leq \phi(x)$ for $x \geq 1$, we have

$$\begin{aligned} \sum_{i=1}^d \mathbb{P}(X(t_i) > a \log(n_0 + k) + C) &\leq d \bar{\Phi}((a \log(n_0 + k) + C)/\bar{\sigma}) \\ &\leq d \phi((a \log(n_0 + k) + C)/\bar{\sigma}) \\ &\leq d \int_{k-1}^k \phi((a \log(n_0 + s) + C)/\bar{\sigma}) ds \\ &= d \int_0^\infty \phi((a \log(n_0 + s) + C)/\bar{\sigma}) ds g_{n_0}(k) \\ &= \frac{de^{-C/a}}{\sqrt{2\pi}} \frac{\bar{\Phi}((a \log n_0 + C)/\bar{\sigma} - \bar{\sigma}/a)}{\phi(\bar{\sigma}/a)/(\bar{\sigma}/a)} g_{n_0}(k), \end{aligned}$$

where the last equality is established in the proof of Lemma 5.2. By assumption, the factor in front of $g_{n_0}(k)$ is bounded by δ . This proves the first claim.

Applying Proposition 5.1 and summing the preceding display over k , we find that the probability that `SAMPLESINGLERECORD` does not return ‘degenerate’ is

$$\begin{aligned} \sum_{k=1}^\infty \mathbb{P}(T_{n_1} = k) &\leq \sum_{k=1}^\infty \sum_{i=1}^d \mathbb{P}(X(t_i) > a \log(n_1 + k) + C) \\ &\leq \sum_{k=1}^\infty \sum_{i=1}^d \mathbb{P}(X(t_i) > a \log(n_0 + k) + C) \\ &\leq \frac{de^{-C/a}}{\sqrt{2\pi}} \frac{\bar{\Phi}((a \log n_0 + C)/\bar{\sigma} - \bar{\sigma}/a)}{\phi(\bar{\sigma}/a)/(\bar{\sigma}/a)} \leq \delta, \end{aligned}$$

which proves the second claim. \square

5.2. Beyond N_X . We next describe how to sample (X_1, \dots, X_n) conditionally on $T_{n_1} = \infty$. As in Section 4.2 we use an acceptance/rejection algorithm, but we have to modify the procedure slightly because we work with a sequence of iid random fields instead of a random walk.

Function `SAMPLEWITHOUTRECORDX` (n_1, ℓ) : **Samples** (X_1, \dots, X_ℓ) **conditionally on** $T_{n_1} = \infty$ **for** $\ell \geq 1$

Step 1: Repeat
 Step 2: $X \leftarrow \text{sample}(X_1, \dots, X_\ell)$ under \mathbb{P}
 Step 3: Until $\sup_{1 \leq k \leq \ell} [X_k - a \log(n_1 + k)] < C$
 Step 4: Return X
 Step 5: EndFunction

5.3. The full algorithm. We summarize our findings in this section in our full algorithm for sampling $(X_1, \dots, X_{N_X+\ell})$ under \mathbb{P} given some $\ell \geq 0$.

The idea is to successively apply SAMPLESINGLERECORD to generate the η_i from the beginning of this section. Starting from $\eta_0 = n_0$ satisfying the requirements in Proposition 5.3, n_1 replaced by each of the subsequent η_i . As a result, we have $\mathbb{P}(\eta_i = \infty | \eta_{i-1} < \infty) \geq 1 - \delta$ by Proposition 5.3. Thus, the number of records is bounded in probability by a geometric random variable with parameter $1 - \delta$.

Algorithm X: Samples $(X_1, \dots, X_{N_X+\ell})$ **given** $a, \delta \in (0, 1)$, $C \in \mathbb{R}$, $\sigma > 0$, $\ell \geq 0$. # n_0 must satisfy the requirements in Proposition 5.3.

Step 1: $X \leftarrow []$, $\eta \leftarrow n_0$
 Step 2: $X \leftarrow \text{sample}(X_1, \dots, X_\eta)$ under \mathbb{P}
 Step 3: Repeat
 Step 4: $\text{segment} \leftarrow \text{SAMPLESINGLERECORD}(a, C, n_0, \eta)$
 Step 5: If segment is not ‘degenerate’
 Step 6: $X \leftarrow [X, \text{segment}]$
 Step 7: $\eta \leftarrow \text{length}(X)$
 Step 8: EndIf
 Step 9: Until segment is ‘degenerate’
 Step 10: If $\ell > 0$
 Step 11: $X \leftarrow [X, \text{SAMPLEWITHOUTRECORDX}(\eta, \ell)]$
 Step 12: EndIf

6. FINAL ALGORITHM AND PROOF OF THEOREM 2.1

In this section, we give our final algorithm. We also give the remaining arguments showing why it produces exact samples and prove a bound on the computational complexity. Together these proofs establish Theorem 2.1.

We start with a description of our final algorithm for sampling M , which exploits that for $S_n = \gamma n - A_n$ and $N_A = N_S$, we have $S_n < 0$ and therefore $A_n \geq \gamma n$ for $n > N_A$.

Algorithm M: Samples $(M(t_1), \dots, M(t_d))$ **given** $\delta, a \in (0, 1)$, $\gamma < \mathbb{E}A_1$, $C \in \mathbb{R}$, σ .

Step 1: Sample A_1, \dots, A_{N_A} using Steps 1–9 from **Algorithm S** with $S_n = \gamma n - A_n$.
 Step 2: Sample X_1, \dots, X_{N_X} using Steps 1–9 from **Algorithm X**.
 Step 3: Calculate N_a with (3.2) and set $N = \max(N_A, N_X, N_a)$.
 Step 4: If $N > N_A$
 Step 5: Sample A_{N_A+1}, \dots, A_N as in Step 10–12 from **Algorithm S** with $S_n = \gamma n - A_n$.
 Step 6: EndIf
 Step 7: If $N > N_X$
 Step 8: Sample X_{N_X+1}, \dots, X_N as in Step 10–12 from **Algorithm X**.
 Step 9: EndIf
 Step 10: Return $M(t_i) = \max_{1 \leq n \leq N} \{-\log A_n + X_n(t_i) + \mu(t_i)\}$ for $i = 1, \dots, d$.

The pathwise construction in Section 3 implied that the output of Algorithm 6 is an exact sample of $\{M(t_1), \dots, M(t_d)\}$. The running time of **Algorithm M** remains to study.

6.1. Computational complexity. We next study the truncation point N in (3.3). Because N_X is bounded in probability by a geometric random variable, it is clear that $N < \infty$ almost surely.

It is our next aim to show how our algorithm depends on d . The only places where d enters the algorithm are in the definition of n_0 and the measure $\mathbb{P}^{(n)}$. Sampling from the latter happens at most a geometric number of times with parameter $1 - \delta$, so the computational complexity is dominated by the choice of n_0 .

For any $\zeta > 0$, if d is large enough and ignoring rounding, the choice $n_0 = n_0(d)$ given by

$$\log(n_0(d)) = \frac{\bar{\sigma}^2}{a^2} - \frac{C}{a} + \frac{\bar{\sigma}}{a} \sqrt{(2 + \zeta) \log \left(\frac{de^{-C/a}}{\delta \sqrt{2\pi} \phi(\bar{\sigma}/a) / (\bar{\sigma}/a)} \right)}$$

satisfies the requirements of Proposition 5.3.

The next lemma is useful to show the second part of Theorem 2.1.

Lemma 6.1. *If K has probability mass function g_{n_0} , then for $\forall p \geq 1$, $\log(\mathbb{E}[K^p]) = O(\log n_0)$.*

Proof. Assume n_0 sufficiently large,

$$\begin{aligned} \mathbb{E}[K^p] &= \sum_{k=1}^{\infty} k^p g_{n_0}(k) \\ &\leq \frac{\int_0^{\infty} (s + n_0)^p \phi((a \log(n_0 + s) + C)/\bar{\sigma}) ds}{\int_0^{\infty} \phi((a \log(n_0 + s) + C)/\bar{\sigma}) ds} \\ &= e^{\frac{p^2 \bar{\sigma}}{2a^2} - \frac{Cp}{a}} \frac{\bar{\Phi}((a \log(n_0) + C - p\bar{\sigma}^2/a)/\bar{\sigma} - \bar{\sigma}/a)}{\bar{\Phi}((a \log(n_0) + C)/\bar{\sigma} - \bar{\sigma}/a)} \\ &\leq e^{\frac{p^2 \bar{\sigma}}{2a^2} - \frac{Cp}{a}} \frac{1}{\frac{(a \log(n_0) + C - p\bar{\sigma}^2/a)/\bar{\sigma} - \bar{\sigma}/a}{((a \log(n_0) + C)/\bar{\sigma} - \bar{\sigma}/a)^2 + 1}} \phi((a \log(n_0) + C - p\bar{\sigma}^2/a)/\bar{\sigma} - \bar{\sigma}/a) \\ &\leq 2e^{\frac{p^2 \bar{\sigma}}{2a^2} - \frac{Cp}{a}} \exp \left(-\frac{p^2 \bar{\sigma}^2}{2a^2} + \frac{p\bar{\sigma}}{a} (a \log(n_0) + C)/\bar{\sigma} - \bar{\sigma}/a \right) \\ &= 2 \exp \left(p \log(n_0) - \frac{p\bar{\sigma}^2}{a^2} \right), \end{aligned}$$

therefore $\log \mathbb{E}[K^p] \leq p \log(n_0) + \log 2 - p\bar{\sigma}^2/a^2$. □

Now we are ready to show that $\log \mathbb{E}[N_X^p] = O(\sqrt{\log d})$. We have

$$N_X = n_0 + \sum_{i=1}^G K_i,$$

where K_i are iid generated from $g_{n_0}(\cdot)$, G is the last time that the segment is not 'degenerate.' Proposition 5.3 shows that G is bounded by a geometric random variable G' with parameter δ almost surely, while G' is independent of the sequence $(K_i)_{i=1}^{\infty}$. Therefore, we have

$$\begin{aligned} \mathbb{E}[N_X^p] &\leq \mathbb{E} \left[\left(n_0 + \sum_{i=1}^{G'} K_i \right)^p \right] \\ &\leq \mathbb{E} \left[\frac{n_0^p + \sum_{i=1}^{G'} K_i^p}{G'} G'^p \right] \\ &= n_0^p \mathbb{E}[G'^{p-1}] + \mathbb{E}[K_i^p] \mathbb{E}[G'^p]. \end{aligned}$$

Therefore we have shown that $\log \mathbb{E}[N_X^p] = O(\sqrt{\log d})$, which means $\mathbb{E}[N_X^p]$ increases slower than d^ϵ for any $\epsilon > 0$.

Clearly N_A or N_a does not depend on d . We only need to show $\mathbb{E}[N_A^p] < \infty$, and $\mathbb{E}[N_a^p] < \infty$.

Recall that in Section 4 we sample the downcrossing segment of the random walk with the nominal distribution, then the upcrossing segment with the exponential tilted distribution. We denote the i 'th downcrossing segment having length τ_i^- , and the i 'th upcrossing segment having length τ_i^+ . Therefore,

$$N_A = \sum_{i=1}^L (\tau_i^- + \tau_i^+),$$

where L is the first time that the upcrossing segment is 'degenerate.' Recall that τ^+ denotes the first upcrossing time of level 0. Because for any $x \leq 0$,

$$\mathbb{P}_x(\tau^+ = \infty) \geq \mathbb{P}_0(\tau^+ = \infty) > 0,$$

L is a.s. bounded by a geometric random variable L' with parameter $q < 1$. under exponential tilted distribution.

According to the discussion in Remark 3.1, we may assume without loss of generality that A_n has step sizes bounded by $r > 0$. Therefore, $S_{\tau_i^+} \leq \gamma$, and $S_{\tau_i^-} \geq r$. Thus, with Theorem 8.1 in [4], for any $p \geq 1$ and $\epsilon > 0$, there exists some constant $V > 0$, such that

$$\mathbb{E}[(\tau_i^-)^{p(1+\epsilon)}] < V \quad \text{and} \quad \mathbb{E}[(\tau_i^+)^{p(1+\epsilon)}] < V.$$

Thus we obtain

$$\begin{aligned} \mathbb{E}[N_A^p] &\leq \mathbb{E} \left[\left(\sum_{i=1}^{L'} (\tau_i^- + \tau_i^+) \right)^p \right] \\ &\leq \mathbb{E} \left[\frac{\sum_{i=1}^{L'} ((\tau_i^-)^p + (\tau_i^+)^p)}{2L'} (2L')^p \right] \\ &\leq \sum_{i=1}^{\infty} \mathbb{E} [((\tau_i^-)^p + (\tau_i^+)^p) \mathbb{1}(L' \geq i) (2L')^{p-1}] \\ &= 2V^{\frac{1}{1+\epsilon}} \left(\mathbb{E} \left[(2L')^{(p-1)\frac{\epsilon}{1+\epsilon}} L' \right] \right)^{\frac{1+\epsilon}{\epsilon}} < \infty. \end{aligned}$$

The value of N_a is only required to satisfy

$$N_a \geq \left(\frac{A_1 \exp(C - \underline{X}_1)}{\gamma} \right)^{\frac{1}{1-a}}.$$

Therefore, we have

$$\mathbb{E}[N_a^p] = \left(\frac{A_1 \exp(C)}{\gamma} \right)^{\frac{p}{1-a}} \mathbb{E}[\exp(-\underline{X}_1)^p]^{\frac{1}{1-a}}.$$

This naturally holds by Assumption B3).

6.2. Choosing a , C , and γ . Although the value of $a \in (0, 1)$ and $\gamma \in (0, \mathbb{E}[A_1])$ and $C \geq 0$ do not affect the order of the computational complexity of our algorithm, we are still interested in discussing some guiding principles which can be used to choose those parameters for a reasonably good implementation.

First, note that among N_X , N_A , and N_a , only N_X would increase as the number of locations sampled, d , increases. Assuming that C has been fixed, we can see that N_X decreases pathwise

while a increases, therefore we should try to choose a close to 1. On the other hand,

$$N_a \geq \left(\frac{A_1 \exp(C - \underline{X}_1)}{\gamma} \right)^{\frac{1}{1-a}}.$$

If $A_1 \exp(C - \underline{X}_1) > \gamma$, then $N_a \nearrow \infty$ while $a \nearrow 1$. This analysis highlights a trade-off between the values of N_X and N_a with respect to the choice of a . Because $E[N_X]$ is not explicitly tractable, we can have a reasonable balancing of the computational effort by equating n_0 with $E[N_a]$. In particular, we look for the largest value of $a \in (0, 1)$ satisfying the following equation

$$(6.1) \quad \exp \left(\frac{\bar{\sigma}}{a} \Phi^{-1} \left(\delta \sqrt{2\pi} \frac{\phi(\bar{\sigma}/a)}{d\bar{\sigma}/a} \right) + \frac{\bar{\sigma}^2}{a^2} - \frac{C}{a} \right) = \mathbb{E} \left[\left(\frac{A_1 \exp(C - \underline{X}_1)}{\gamma} \right)^{\frac{1}{1-a}} \right];$$

Note that the left-hand side converges to infinity as $a \searrow 0$ while the right-hand side is bounded, but the right-hand side converges to infinity as $a \nearrow 1$ while the left-hand side is bounded, so a solution exists. Such a solution can be obtained by running a pilot run of \underline{X}_1 , then search for the desired a numerically.

Another approach consists in selecting $a = 1$ and adjusting C so that (3.2) holds true for all $n \geq 1$. Therefore, we choose

$$C = \underline{X}_1 + \log \left(\frac{A_1}{\gamma} \right).$$

The value of C is random, but the algorithms can be modified accordingly, modifying the definition of n_0 , which depends on C . However, the expected computational cost remains the same order as the case when C is deterministic as defined earlier.

Similarly, N_A increases pathwise while γ increases, while N_a decreases if γ increases. One could get the empirical average value of N_A via simulation, and choose γ accordingly such that N_A and N_a are balanced.

7. TOLERANCE ENFORCED SIMULATION

In this section we illustrate a general procedure which can be applied so that, for any given $\delta > 0$ one can construct a fully simulatable process $M_\delta(\cdot)$, with the property that

$$\mathbb{P} \left(\sup_{t \in T} |M(t) - M_\delta(t)| \leq \delta \right) = 1.$$

We concentrate on the case in which $T = [0, 1]$ to facilitate the notation. The technique can be easily adapted to higher-dimensional sets T , but the important assumption involves an infinite series representation for X which satisfies certain regularity conditions as we shall explain.

The technique presented in this section is not limited to Gaussian processes, and we do not make this assumption here. As a result, we do not use Assumptions B1)-B3) in this section, but we replace them with A)-D) below. However, Assumptions A1) and A2) on the renewal sequence (A_n) are in force throughout this section.

7.1. An infinite series representation. We assume that $(X_n(t))_{t \in T}$ can be expressed as an almost surely convergent series of basis functions with random weights. We illustrate the procedure with a particularly convenient family of basis functions.

First, let us write any $m \geq 1$ as $m = 2^j + k$ for $j \geq 0$ and $0 \leq k \leq 2^j - 1$, and note that there is only one way to write m in this form. We assume that there exists a sequence of basis functions $(\Lambda_m(\cdot))_{m \geq 0}$, with support on $[0, 1]$ (i.e., $\Lambda_m(t) = 0$ for $t \notin [0, 1]$). Moreover, we assume that $|\Lambda_0(t)|, |\Lambda_1(t)| \leq 1$ for all $t \in [0, 1]$, and that for every $m \geq 1$,

$$\Lambda_m(t) = \Lambda_1(2^j(t - k/2^j)).$$

We introduce normalizing constants, $\lambda_0 > 0$ and $\lambda_m = \lambda' 2^{-j\alpha}$ for $m \geq 1$, where $\alpha \in (0, 1)$ and $\lambda' > 0$. Finally, we assume that

$$X_n(t) = \sum_{m=0}^{\infty} Z_{m,n} \Lambda_m(t) \lambda_m,$$

where the random variables $(Z_{m,n})_{m \geq 0, n \geq 1}$ are iid. We shall use Z to denote a generic copy of the $Z_{m,n}$'s and we shall impose suitable assumptions on the tail decay of Z . The parameter α relates to the Hölder continuity exponent of the process X_n . For example, if X_n is Brownian motion, $\alpha = 1/2$. This interpretation of α will not be used in our development, but it helps to provide intuition which can be used to inform the construction of a model based on the basis functions that we consider. For more information on the connection to the Hölder properties implied by α , the reader should consult [9] and the references therein.

Throughout, we use the following total order among the pairs $\{(m, n) : m \geq 0, n \geq 1\}$. We say $(m, n) < (m', n')$ if $m + n < m' + n'$ and in case $m + n = m' + n'$, we say that (m, n) is smaller than (m', n') in lexicographic order. In particular, we have

$$(0, 1) < (0, 2) < (1, 1) < (0, 3) < (1, 2) < (2, 1) < \dots$$

We let $\theta(m, n)$ be the position of (m, n) in the total order. We also define $\eta(\cdot) : \mathbb{N} \rightarrow \mathbb{N} \cup \{0\} \times \mathbb{N}$ to be the inverse function of $\theta(\cdot)$, and given $\theta \in \mathbb{N}$, we write

$$\eta(\theta) = (\eta_m(\theta), \eta_n(\theta)).$$

7.2. Building blocks for our algorithm. We now proceed to describe the construction of M_δ , which is adapted from a record-breaking technique introduced in [7]. To define M_δ , we use the truncated series

$$X_n(t; K) = \sum_{m \leq K} \lambda_m Z_{m,n} \Lambda_m(t).$$

We abuse notation by re-using notation such as N_X and N_A throughout our discussion of TES, but the random variables are not the same as in the rest of the paper.

Our algorithm relies on three random times to be defined next. The first is a positive integer-valued random variable N_X and functions $a(\cdot)$, $\xi_0(\cdot)$, $\xi_1(\cdot)$, which satisfy, for $k \geq N_X$ and $n \geq 1$,

$$\sup_{t \in T} |X_n(t) - X_n(t; k)| \leq \xi_1(k) + \xi_0(k) a(n)$$

and, for $n \geq N_X$,

$$\sup_{t \in T} |X_n(t)| \leq (a(0)\lambda_0 + \xi_1(1)) + (\lambda_0 + \xi_0(1))a(n),$$

see Proposition 7.2 below. The second is $N_A = N_A(\gamma)$, which satisfies, for $n \geq N_A$,

$$A_n \geq \gamma n,$$

and we can sample N_A jointly with (A_1, \dots, A_{N_A}) using **Algorithm S** in Section 4. The third is N_ξ , which satisfies, for $n \geq N_\xi$,

$$(7.1) \quad (a(0)\lambda_0 + \xi_1(1)) + (\lambda_0 + \xi_0(1))a(n) - \log(n\gamma) \leq \inf_{t \in [0,1]} X_1(t, N_X) - \log(A_1) - \xi_1(N_X) - \xi_0(N_X)a(n).$$

Setting $N = \max(N_X, N_A, N_\xi)$, we have, for $t \in T$ and $n \geq N$,

$$\begin{aligned} -\log(A_n) + X_n(t) &\leq -\log A_n + (a(0)\lambda_0 + \xi_1(1)) + (\lambda_0 + \xi_0(1))a(n) \\ &\leq -\log(n\gamma) + (a(0)\lambda_0 + \xi_1(1)) + (\lambda_0 + \xi_0(1))a(n) \\ &\leq -\log(A_1) + \inf_{t \in [0,1]} X_1(t, N_X) - \xi_1(N_X) - \xi_0(N_X)a(n) \\ &\leq -\log(A_1) + \inf_{t \in [0,1]} X_1(t) \end{aligned}$$

$$\leq -\log(A_1) + X_1(t),$$

and therefore, for $t \in T$,

$$(7.2) \quad \sup_{n \geq 1} \{-\log A_n + X_n(t) + \mu(t)\} = \max_{1 \leq n \leq N} \{-\log A_n + X_n(t) + \mu(t)\}.$$

If we select an integer $K_\delta \geq N_X$ such that $\xi_1(K_\delta) + \xi_0(K_\delta)a(n) \leq \delta$, then

$$M_\delta(t) = \max_{1 \leq n \leq N} \{-\log A_n + X_n(t; K_\delta) + \mu(t)\}$$

satisfies $\sup_{t \in T} |M(t) - M_\delta(t)| \leq \delta$.

It remains to explain how to simulate N_X jointly with (X_1, \dots, X_N) and how to construct ξ_0 , and ξ_1 . For this, we use a variant of the record-breaking technique, but we first need to discuss our assumptions on the $Z_{m,n}$'s.

7.3. Assumptions on the $Z_{m,n}$'s and an example. We introduce some assumptions on the distribution of Z in order to use our record-breaking algorithm. We write $\bar{F}(\cdot)$ for the tail distribution of $|Z|$, that is $\bar{F}(t) = \mathbb{P}(|Z| > t)$ for $t \geq 0$. Assume that we can find: a bounded and nonincreasing function $\bar{H}(\cdot)$ on $[0, \infty)$, an easy-to-evaluate eventually nonincreasing function $\Gamma(\cdot)$ on \mathbb{N} , as well as some $\theta_0 > 0$, $b \in (0, 1)$, and $\rho > 0$ satisfying the following assumptions with

$$a(n) = \rho(\log(n+1))^b :$$

- A) For (m, n) satisfying $\theta(m, n) \geq \theta_0$, we have $\bar{F}(a(m) + a(n)) \leq \bar{H}(a(m))\bar{H}(a(n))$.
- B) We have $\sum_{m=0}^{\infty} \bar{H}(a(m)) < \infty$.
- C) For $r > \theta_0$, we have $1 > \Gamma(r) \geq \sum_{(m,n): \theta(m,n) > r} \bar{H}(a(m))\bar{H}(a(n))$.
- D) We have $\sum_r r^\varepsilon \Gamma(r) < \infty$ for some $\varepsilon > 0$.

Assumptions A), B), and C) are needed to run the algorithm, and Assumption D) to bound moments of the computational complexity.

As an example, we now show that these assumptions are satisfied when $X_n(\cdot)$ is a Brownian motion. Similar constructions are possible for fractional Brownian motion [6], but we do not work out the details here. First, $\Lambda_0(t) = tI(t \in [0, 1])$, $\Lambda_1(t) = 2tI(t \in [0, 1/2]) + 2(1-t)I(t \in (1/2, 1])$, $\alpha = 1/2$, and $\lambda_0 = \lambda' = 1$; see [30]. Second, the $Z_{m,n}$'s are iid standard Gaussian random variables and one can select $\bar{H}(t) = \phi(t)$, the standard normal density, so that we have Assumption A) for $\theta_0 = \inf\{\theta : a(\eta_m(\theta)) + a(\eta_n(\theta)) \geq 2\sqrt{2\pi}\}$ and B) is evident. Moreover, selecting any $\rho > 4$ and $b = 1/2$ allows us to satisfy Assumptions C) and D). Indeed, note that

$$\begin{aligned} & \sum_{\theta(m,n) \geq r} \bar{H}(a(m))\bar{H}(a(n)) \\ &= \sum_{\theta(m,n) \geq r} \left(\frac{2}{\pi}\right) \exp\left(-\rho^2 \frac{\log(m+1) + \log(n+1)}{2}\right) \\ &= \sum_{\theta(m,n) \geq r} \left(\frac{1}{(m+1)(n+1)}\right)^{\rho^2/2} \leq \sum_{\theta(m,n) \geq r} \left(\frac{1}{m+n}\right)^{\rho^2/2}. \end{aligned}$$

The point (m, n) with $\theta(m, n) = r$ is one of the $\ell(r)$ points on the segment between $(\ell(r), 0)$ and $(1, \ell(r) - 1)$, where $\ell(r) = \lceil \sqrt{2r + 1/4} - 1/2 \rceil$. We therefore continue to bound as follows:

$$\sum_{k \geq \ell(r)} k^{1-\rho^2/2} \leq \int_{\ell(r)-1}^{\infty} x^{1-\rho^2/2} dx = \frac{1}{\rho^2/2 - 2} (\ell(r) - 1)^{2-\rho^2/2}.$$

Thus, in the Brownian case we can define $\Gamma(r)$ to be the right-hand side of the preceding display, so for instance any $\rho > 4$ implies Assumption D).

Let us discuss the implications of our strategy when applied to the Brownian case. In the case in which $X_n(\cdot)$ is a standard Brownian motion we have

$$\sum_{m=0}^{2^r-1} \lambda_m Z_{m,n} \Lambda_m(t) = X_n(t),$$

for every dyadic point $t = j2^{-r}$ with $j = 0, 1, \dots, 2^r$. Therefore, once we fix any $\delta > 0$ (say $\delta = 1/2$), we can apply the previous strategy to obtain N and we can continue sampling $Z_{m,n}$ for $m \geq K_\delta$ if needed so that we can return

$$M(t) = \max_{1 \leq n \leq N} \left\{ -\log(A_n) + \sum_{m=0}^{2^r-1} \lambda_m Z_{m,n} \Lambda_m(t) \right\}.$$

Consequently, we conclude that at least in the Brownian case the procedure that we present here can be used to evaluate $\{M(j/2^r)\}_{j=0}^d$ with $d = 2^r$ exactly and with expected computational cost of order $O(d \cdot \mathbb{E}[N]) = O(d)$ – because $\mathbb{E}[N]$ does not depend on d and, as we show in Theorem 7.4 below, $\mathbb{E}[N] < \infty$.

7.4. Breaking records for the $Z_{m,n}$'s. Define $T_0 = 0$, and, for $k \geq 1$,

$$T_k = \inf\{\theta(m, n) > T_{k-1} : |Z_{m,n}| > a(m) + a(n)\}.$$

In this subsection, given some integer $\theta_0 \geq 0$, we develop a technique to sample the random set $\mathcal{T} = \{T_k : T_k < \infty\} \cap \{\theta_0 + 1, \dots\}$ jointly with $(Z_{m,n})_{m \geq 0, n \geq 1}$. Indeed, given \mathcal{T} , the $Z_{m,n}$ are independent and have the following distributions. For $\theta(m, n) \leq \theta_0$, $Z_{m,n}$ has the nominal (unconditional) distribution. For $\theta(m, n) \in \mathcal{T}$, $Z_{m,n}$ has the conditional distribution of Z given $\{|Z| > a(m) + a(n)\}$, and if $\theta(m, n) \notin \mathcal{T}$, $Z_{m,n}$ has the conditional distribution of Z given $\{|Z| \leq a(m) + a(n)\}$.

We first note that that only finitely many T_k 's are finite, so that we can once again apply a record breaking technique, based on the record-breaking epochs T_k . Indeed, applying Assumptions A) and B), we find that

$$\sum_{m,n} P(|Z_{m,n}| > a(m) + a(n)) \leq \sum_{m,n} \bar{H}(a(m)) \bar{H}(a(n)) = \left(\sum_m \bar{H}(a(m)) \right)^2 < \infty,$$

and the claim follows from the Borel-Cantelli lemma.

The function `SAMPLERECORDSZ` given below, which is directly adapted from Algorithm 2w in [7], allows one to sequentially sample the elements in $\{T_k : T_k < \infty\}$ jointly with the $Z_{m,n}$'s. The function `SAMPLERECORDSZ` takes as input θ_0 satisfying $\Gamma(\theta_0) < 1$.

Function `SAMPLERECORDSZ`(θ_0): Samples the set $\mathcal{T} = \{T_k : T_k < \infty\} \cap \{\theta_0 + 1, \dots\}$

Step 1: Initialize $G \leftarrow \theta_0$ and $\mathcal{T} \leftarrow []$.

Step 2: $u \leftarrow 1$, $d \leftarrow 0$. $V \leftarrow U(0, 1)$.

Step 3: While $u > V > d$

Step 4: $G \leftarrow G + 1$

Step 5: $d \leftarrow \max(d, (1 - \Gamma(G)) \times u)$

Step 6: $u \leftarrow \mathbb{P}(|Z| \leq a(\eta_m(G)) + a(\eta_n(G))) \times u$

Step 7: EndWhile

Step 8: If $V \geq u$, then $\mathcal{T} \leftarrow [\mathcal{T}, G]$ and go to Step 2.

Step 9: If $V \leq d$, stop and return \mathcal{T} .

The next proposition establishes that the output of the function `SAMPLERECORDSZ` has the desired distribution.

Proposition 7.1. *The output from $\text{SAMPLERECORDSZ}(\theta_0)$ is a sample of the set $\mathcal{T} = \{T_k : T_k < \infty\} \cap \{\theta_0 + 1, \dots\}$. Moreover, we have $\mathbb{E}[(\max(0, \sup \mathcal{T}))^\beta] < \infty$ for some $\beta > 1$.*

Proof. For simplicity we assume throughout this proof that $\theta_0 = 0$. For the first claim it suffices to show that $\text{SAMPLERECORDSZ}(0)$ returns $\mathcal{T} = \inf\{T_k : T_k < \infty\} \cap \{1, 2, \dots\}$ without bias; subsequent T_k 's follow by the same reasoning. We write $T = T_1$.

In Steps 3 through 5 the algorithm iteratively constructs the sequences (u_j) and (d_j) given by

$$u_j = u_{j-1} \mathbb{P}(|Z| \leq a(\eta_m(j)) + a(\eta_n(j))), \quad d_j = \max(d_{j-1}, u_{j-1}(1 - \Gamma(j)))$$

with $u_0 = 1$ and $d_0 = 0$. It is evident that both sequences are monotone. Moreover, we have $u_j = \mathbb{P}(T > j)$ for $j \geq 0$ and $\lim_{j \rightarrow \infty} u_j = \mathbb{P}(T = \infty)$. Similarly, because $\lim_{j \rightarrow \infty} \Gamma(j) = 0$ we obtain $\lim_{j \rightarrow \infty} d_j = \mathbb{P}(T = \infty)$.

Let $n(V)$ be the number of times Step 3 is executed before either going to Step 8 or Step 9. It suffices to check that when Step 8 is executed then the element added to \mathcal{T} has the law of T given $T < \infty$, and that Step 9 is executed with probability $\mathbb{P}(T = \infty)$. For the former, we note that by definition of $n(V)$ and because $u_j \in (d_{j-1}, u_j)$, we have for $j \geq 1$

$$\begin{aligned} \mathbb{P}(n(V) = j \mid V \geq u_{n(V)}) &= \frac{\mathbb{P}(V \in (d_{j-1}, u_{j-1}), V \geq u_j)}{\mathbb{P}(V \geq u_{n(V)})} \\ &= \frac{\mathbb{P}(V \in (u_j, u_{j-1}))}{\mathbb{P}(V \geq u_{n(V)})} \\ &= \frac{u_{j-1} - u_j}{1 - \lim_{k \rightarrow \infty} u_k}, \end{aligned}$$

which equals $\mathbb{P}(T = j \mid T < \infty)$ as desired. For the latter, we note that

$$\mathbb{P}(V \leq d_{n(V)}) = 1 - \mathbb{P}(V \geq u_{n(V)}) = \mathbb{P}(T = \infty).$$

In preparation for the proof of the second claim of the proposition, we bound the probability that the while loop requires more than $k \geq 1$ iterations:

$$\begin{aligned} \mathbb{P}(n(V) > k) &\leq \mathbb{P}(V \in (d_k, u_k)) \leq \mathbb{P}(V \in (d_k, u_{k-1})) = u_{k-1} - d_k \\ &= u_{k-1} - \max\{d_{k-1}, u_{k-1}(1 - \Gamma(k))\} \leq \Gamma(k). \end{aligned}$$

As a consequence of the inequality

$$\mathbb{E}[n(V)^\beta] = \sum_{k=0}^{\infty} ((k+1)^\beta - k^\beta) \mathbb{P}(n(V) > k) \leq 1 + \sum_{k=1}^{\infty} ((k+1)^\beta - k^\beta) \Gamma(k),$$

we find that $\mathbb{E}[n(V)^\beta] < \infty$ if $\sum_k k^{\beta-1} \Gamma(k) < \infty$.

We have a similar finite-moment bound for subsequent calls to the while loop. Writing $n_i(V_1, \dots, V_i)$ for the number of iterations in the i -th execution of the while loop, where V_1, V_2, \dots are the iid standard uniform random variables generated in subsequent calls to Step 2. Compared to the above argument for $i = 1$, this quantity only depends on V_1, \dots, V_{i-1} through a random shift of Γ . Because Γ is eventually nonincreasing, there exists a constant c' such that, for all $i \geq 1$,

$$(7.3) \quad \mathbb{E}[n_i(V_1, \dots, V_i) \mid V_1, \dots, V_{i-1}] \leq c' \sum_k k^{\beta-1} \Gamma(k).$$

To prove a bound on the moment of $\sup \mathcal{T}$, we first let Υ be the number of times we execute the while loop. We then note that, for any random variable G and any $\beta \geq 1$, by Jensen's inequality,

$$\max(0, \sup \mathcal{T})^\beta = \left(\sum_{i=1}^{\Upsilon-1} n_i(V_1, \dots, V_i) \right)^\beta = \left(\sum_{i=1}^{\infty} n_i(V_1, \dots, V_i) I(\Upsilon > i) \right)^\beta$$

$$\begin{aligned}
&\leq \sum_{i=1}^{\infty} \left(\frac{n_i(V_1, \dots, V_i) I(\Upsilon > i)}{\mathbb{P}(G = i)} \right)^{\beta} \mathbb{P}(G = i) \\
&= \sum_{i=1}^{\infty} n_i(V_1, \dots, V_i)^{\beta} I(\Upsilon > i) \mathbb{P}(G = i)^{1-\beta},
\end{aligned}$$

because the right-hand side is finite almost surely.

Because the event $\{\Upsilon > i - 1\}$ only depends on V_1, \dots, V_{i-1} , we have by (7.3),

$$\begin{aligned}
\mathbb{E}[n_i(V_1, \dots, V_i)^{\beta}, \Upsilon > i] &\leq \mathbb{E}[n_i(V_1, \dots, V_i)^{\beta}, \Upsilon > i - 1] \\
&= \mathbb{E}[I(\Upsilon > i - 1) \mathbb{E}[n_i(V_1, \dots, V_i) | V_1, \dots, V_{i-1}]] \\
&\leq c' \left(\sum_k k^{\beta-1} \Gamma(k) \right) P(\Upsilon > i - 1) \\
&\leq c' \left(\sum_k k^{\beta-1} \Gamma(k) \right) P(T_1 < \infty)^{i-1},
\end{aligned}$$

where we use the fact that Υ is stochastically dominated by a geometric random variable with success parameter $P(T_1 = \infty) > 0$. Combining the preceding displays, we deduce that, for $\beta \geq 1$,

$$\mathbb{E} \left[\max(0, \sup \mathcal{T})^{\beta} \right] \leq c' \left(\sum_k k^{\beta-1} \Gamma(k) \right) \sum_{i=1}^{\infty} P(T_1 < \infty)^{i-1} \mathbb{P}(G = i)^{1-\beta},$$

which is seen to be finite for some $\beta > 1$ by Assumption D) upon choosing G geometric with a suitably chosen success probability. \square

7.5. Truncation error of the infinite series. We next write, for $k \geq 0$

$$X_n(t) = X_n(t; k) + \sum_{m > k} \lambda_m Z_{m,n} \Lambda_m(t).$$

and it is our objective to study the truncation error, i.e., the second term.

The next proposition controls the truncation error in terms of functions ξ_0 and ξ_1 defined for $r \geq 1$ through

$$\begin{aligned}
\xi_0(r) &= \lambda'(1 - 2^{-\alpha})^{-1} 2^{-\alpha \lfloor \log_2(r) \rfloor}, \\
\xi_1(r) &= \frac{\rho}{\log_2(e)} \left(\lfloor \log_2(r) \rfloor + \frac{2^{-\alpha}}{1 - 2^{-\alpha}} + 2 \right) \xi_0(r).
\end{aligned}$$

Note that $\xi_0(r), \xi_1(r) \rightarrow 0$ as $r \rightarrow \infty$. We also write

$$N_X = \max\{\sup \mathcal{T}, \theta_0 - 1\}.$$

If \mathcal{T} is empty then $\sup \mathcal{T} = -\infty$ and therefore $N_X = \theta_0 - 1$; otherwise, if \mathcal{T} is non-empty, then $\sup \mathcal{T} \geq \theta_0$ and therefore $N_X \geq \theta_0$.

Proposition 7.2. *For all $K \geq N_X$ and $n \geq 1$, we have*

$$\sup_{t \in T} |X_n(t) - X_n(t; K)| \leq \xi_1(K) + \xi_0(K) a(n).$$

Moreover, for $n \geq N_X$, we have

$$\sup_{t \in T} |X_n(t)| \leq (a(0)\lambda_0 + \xi_1(1)) + (\lambda_0 + \xi_0(1))a(n).$$

Proof. To establish the first inequality, we write

$$|X_n(t) - X_n(t; K)| \leq \sum_{m>K} \lambda_m a(m) |\Lambda_m(t)| + a(n) \sum_{m>K} \lambda_m |\Lambda_m(t)|.$$

If $m > K \geq N_X$, because $\theta(m, n) \geq m$, we have from the definition of N_X , that

$$|\lambda_m Z_{m,n} \Lambda_m(t)| \leq \lambda_m (a(m) + a(n)) |\Lambda_m(t)|.$$

We bound the summand of the second sum by noting that, for $r \geq 1$,

$$\sup_{t \in T} \sum_{m=r}^{\infty} \lambda_m |\Lambda_m(t)| \leq \sup_{t \in T} \sum_{j=\lfloor \log_2(r) \rfloor}^{\infty} \sum_{k=0}^{2^j-1} \lambda_{2^j+k} |\Lambda_{2^j+k}(t)| \leq \sum_{j=\lfloor \log_2(r) \rfloor}^{\infty} \lambda' 2^{-\alpha j} = \xi_0(r).$$

We similarly bound the summand in the first sum, using the definition of $a(\cdot)$ and the fact that

$$\sum_{j=k}^{\infty} j s^j = s^k \frac{k(1-s) + s}{(1-s)^2}$$

for $|s| < 1$.

For the second inequality in the proposition, we note that for $n \geq N_X$,

$$|X_n(t)| \leq \sum_{m=0}^{\infty} |\lambda_m Z_{m,n} \Lambda_m(t)| \leq (a(0) + a(n)) \lambda_0 + \sum_{m=1}^{\infty} \lambda_m (a(m) + a(n)) |\Lambda_m(t)|,$$

because $\theta(m, n) \geq N_X$ for each $n \geq N_X$. The sum over m is bounded by $\xi_1(1) + \xi_0(1)a(n)$ as shown in the proof of the first inequality. \square

7.6. Construction of M_δ . Now we are ready to provide the final algorithm for computing M_δ .

Algorithm TES: Samples M_δ given $\delta > 0$.

- Step 1: $\mathcal{T} \leftarrow \text{Sample SAMPLERECORDSZ}(\theta_0)$
- Step 2: $N_X \leftarrow \max\{\sup \mathcal{T}, \theta_0 - 1\}$
- Step 3: Sample $Z_{m,n}$ from the nominal distribution if $\theta(m, n) \leq \theta_0$
- Step 4: For $0 \leq m \leq N_X$ and $\theta(m, 1) > \theta_0$
- Step 5: If $\theta(m, 1) \in \mathcal{T}$: sample $Z_{m,1}$ from the law of Z given $\{|Z| > a(m) + a(1)\}$
- Step 6: Else If: sample $Z_{m,1}$ from the law of Z given $\{|Z| \leq a(m) + a(1)\}$
- Step 7: EndFor
- Step 8: Sample A_1, \dots, A_{N_A} using Steps 1–8 from **Algorithm S** with $S_n = \gamma n - A_n$.
- Step 9: Compute N_ξ , the smallest n for which (7.1) holds, and let $N \leftarrow \max(N_X, N_A, N_\xi)$
- Step 10: Sample A_{N_A+1}, \dots, A_N as in Step 10 from **Algorithm S** with $S_n = \gamma n - A_n$.
- Step 11: Compute the smallest $K_\delta \geq N_X$ such that $\xi_1(K_\delta) + \xi_0(K_\delta)a(N) \leq \delta$.
- Step 12: For $2 \leq n \leq N$, $0 \leq m \leq K_\delta$, $\theta(m, n) > \theta_0$ OR $n = 1$, $N_X < m \leq K_\delta$, $\theta(m, n) > \theta_0$
- Step 13: If $\theta(m, n) \in \mathcal{T}$: sample $Z_{m,n}$ from the law of Z given $\{|Z| > a(m) + a(n)\}$
- Step 14: Else If: sample $Z_{m,n}$ from the law of Z given $\{|Z| \leq a(m) + a(n)\}$
- Step 15: EndFor
- Step 16: Return $M_\delta(t) = \max\{X_n(t; K_\delta) - \log(A_n)\}$.

7.7. Exponential moments of $\sup_{t \in [0,1]} |X(t)|$. We need a bound on the exponential moments of $\sup_{t \in [0,1]} |X(t)|$ in order to analyze N_ξ . If X is Gaussian and continuous, then such a bound immediately follows from Borell's inequality [1, Thm. 2.1.1]. The following proposition establishes the existence of exponential moments in the generality of the present section.

Proposition 7.3. *For any $p > 0$, we have*

$$\mathbb{E} \exp \left(p \sup_{t \in [0,1]} |X(t)| \right) < \infty.$$

Proof. We first note that

$$\sup_{t \in [0,1]} |X_n(t)| \leq \lambda_0 Z_{0,n} + \sum_{j=1}^{\infty} \lambda' 2^{-\alpha j} \max_{k=0, \dots, 2^j-1} |Z_{2^j+k,n}|.$$

It suffices to prove that the tail of the infinite sum in this expression is ultimately lighter than any exponential. A union bound leads to, for $y \geq 0$,

$$\begin{aligned} & \mathbb{P} \left(\sum_{j=1}^{\infty} \lambda' 2^{-\alpha j} \max_{k=0, \dots, 2^j-1} |Z_{2^j+k,n}| > y \right) \\ & \leq \sum_{j=1}^{\infty} \mathbb{P} \left(\lambda' 2^{-\alpha j} \max_{k=0, \dots, 2^j-1} |Z_{2^j+k,n}| > (2^{\alpha/2} - 1) 2^{-\alpha j/2} y \right) \\ & \leq \sum_{j=1}^{\infty} \mathbb{P} \left(\max_{k=0, \dots, 2^j-1} |Z_{2^j+k,n}| > \frac{(2^{\alpha/2} - 1) 2^{\alpha j/2}}{\lambda'} y \right). \end{aligned}$$

Assumptions A) and B) together with Markov's inequality imply that $C' := \mathbb{E} \exp(|Z/\rho|^{1/b}) < \infty$ and therefore, for $t \geq 0$,

$$\mathbb{P} \left(\max_{k=0, \dots, 2^j-1} |Z_{2^j+k,n}| > 2^{\alpha j/2} t \right) \leq 2^j \mathbb{P}(|Z| > 2^{\alpha j/2} t) \leq C' 2^j e^{-(t 2^{\alpha j/2}/\rho)^{1/b}}.$$

Select some $t_0 > 0$ and $\kappa \in (1, 1/b)$ such that $(t 2^{\alpha j/2}/\rho)^{1/b} \geq j + t^\kappa$ for all $j \geq 1$ and $t \geq t_0$. Using this bound results in a tail estimate that is summable over j and lighter than any exponential distribution. \square

7.8. Complexity analysis. We conclude this section with the following result which summarizes the performance guarantee of Algorithm TES. Stronger moment conditions on the computational costs are readily found using the same arguments and a stronger version of Assumption D).

Theorem 7.4. *Suppose that Assumptions A) to D) are in force, together with Assumptions A1) and A2). Given $\delta \in (0, 1)$, the output $(M_\delta(t))_{t \in T}$ of **Algorithm TES** satisfies*

$$\sup_{t \in T} |M_\delta(t) - M(t)| \leq \delta.$$

Moreover, we have

$$\mathbb{E}[K_\delta] = O \left((\delta / \log(1/\delta))^{-1/\alpha} \right),$$

where α is determined by the series representation of X . Finally, the total computational costs of running Algorithm TES has finite expectation and is at most $O((\delta / \log(1/\delta))^{-1/\alpha})$.

Proof. The first claim follows by construction, see Section 7.2.

We have $\mathbb{E}[N_X^\beta] < \infty$ for some $\beta > 1$ from Proposition 7.1. In order to analyze N_ξ , we use Proposition 7.3. Using this estimate it follows from simple calculations and Assumption A1) that $\mathbb{E}[N_\xi^p] < \infty$ for every $p > 0$. We have argued that $\mathbb{E}[N_A^p] < \infty$ in Section 6, so we conclude that $\mathbb{E}[N^\beta] < \infty$. Finally, using the definition of $\xi_0(r)$ and $\xi_1(r)$ we can see that there is a constant $\kappa > 0$ such that

$$K_\delta = O \left(\left[\frac{\delta}{(\log N)^b + \kappa \log(1/\delta)} \right]^{-1/\alpha} \right).$$

This leads to the bound on the first moment of K_δ . The expected running time of the algorithm is order $\mathbb{E}[K_\delta \times N]$, which has a finite first moment because $\mathbb{E}[N^\beta] < \infty$. The complexity bound follows. \square

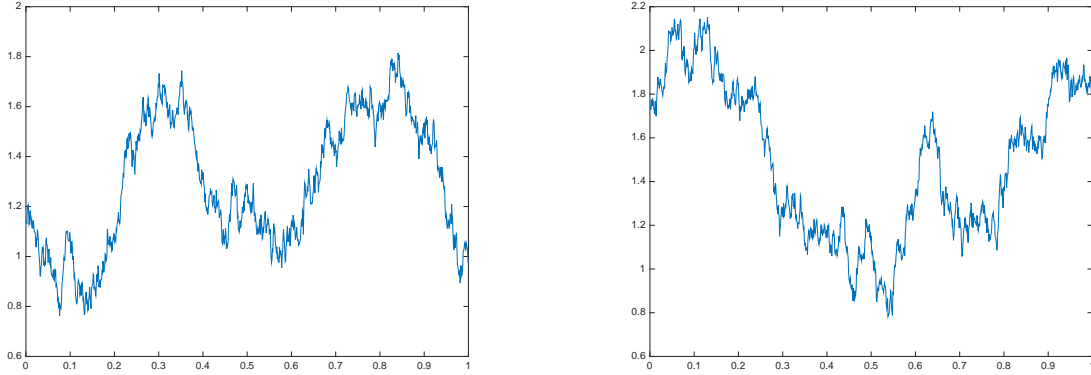


FIGURE 1. The Brown-Resnick process on $[0, 1]$ with Brownian motion input. The grid mesh is 0.001.

8. NUMERICAL RESULTS

In this section we show some simulation results to empirically validate **Algorithm M**. We also compare numerically the computational cost of our method with the exact sampling algorithm proposed in [17]. We implemented our algorithm in Matlab. We chose the value of a and C according to our discussion in Section 6.2. We let $C = 0$, then chose the largest $a \in (0, 1)$ such that 6.1 holds.

We generated the Brown-Resnick processes, $M(t) = \sup_{n \geq 1} \{-\log A_n + X_n(t) - \sigma^2(t)/2\}$, on compact sets. If X is a Brownian motion it was shown in [15] that $M(\cdot)$ has a stationary sample path on $[0, 1]$. Figures 1 shows sample paths of $M(\cdot)$ in this case. Figure 2 presents two samples of the Brown-Resnick random field on $[0, 1]^2$ when X is a Brownian sheet. We validated the implementation of our algorithm by comparing the marginal distribution of our sample with that of the standard Gumbel distribution given by $M(0.5)$. We generated 1000 samples of the Brown-Resnick process M on the dyadic points $\{k/2^5 : 0 \leq k \leq 2^5\}$. The QQ-plot in Figure 3 confirms empirically that the marginal distribution of M is indeed standard Gumbel.

Next we will compare the computational cost of our algorithm and the algorithm proposed in [17]. We conducted both algorithms to generate 200 samples of the Brown-Resnick process $M(\cdot)$ with fractional Brownian motion inputs. We recorded both the mean running time for generating a single sample and the 95% confidence interval based on our 200 samples, for different numbers of grids and with different Hurst parameter of the fractional Brownian motion. We used the grid numbers $d = 1000, 2000, 5000$ and 10000 , and the Hurst parameter $H \in \{1/4, 1/2, 3/4\}$. Figure 4, 5 and 6 show the numerical results, with the solid lines representing the mean running time and the

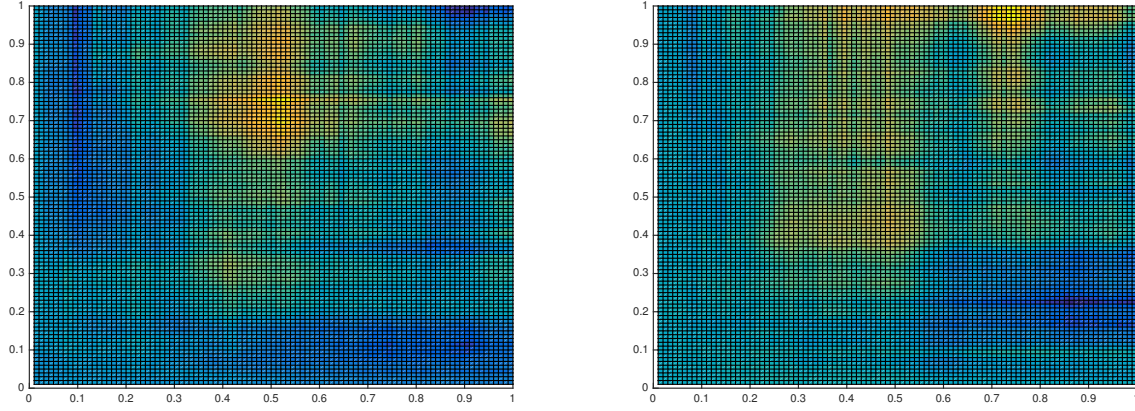


FIGURE 2. The Brown-Resnick field on $[0, 1]^2$ with Brownian sheet input. The grid mesh is 0.001.

dashed lines representing the range of confidence intervals. They illustrate that when the number of grids increases, the computational cost of our algorithm appears to increase almost linearly, while the cost for the algorithm proposed in [17] increases quadratically. Because we are using the circulant embedding method to generate the fractional Brownian vectors, it's consistent with what we expect. One thing to notice is that with such method, the computational cost to generate a d -dimensional Gaussian vector is the same as to generate a $2^{\lceil \log_2 d \rceil}$ -dimensional Gaussian vector. However, this consideration won't affect our comparison because we used this method in both algorithms.

We also counted the number of Gaussian vectors generated in our algorithm and the algorithm of [17]. Figure 7 illustrates our main result that this number increases linearly in the algorithm of [17], while in our algorithm it remains around the same level.

REFERENCES

- [1] Adler, R. J., Taylor, J. E. (2007). *Random Fields and Geometry*. Springer, New York.
- [2] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd ed. Springer, New York.
- [3] Asmussen, S., Glynn, P. (2007). *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.
- [4] Gut, A. (2009). *Stopped Random Walks*. Springer, New York.
- [5] Asmussen, S., Glynn, P. and Pitman, J. (1995). Discretization error in simulation of one-dimensional reflecting Brownian motion. *The Annals of Applied Probability*, 5(4), 875–896.
- [6] Ayache, A. and Taqqu, M.S. (2003) Rate optimality of wavelet series approximations of fractional Brownian motion. *The Journal of Fourier Analysis and Applications*, 9:451–471.
- [7] Blanchet, J., Chen, X. (2015). Steady-state simulation of reflected Brownian motion and related stochastic networks. *The Annals of Applied Probability*, 25(6), 3209–3250.
- [8] Blanchet, J., Chen, X. (2016). Perfect sampling of generalized Jackson networks. *arXiv:1601.05499 [math]*. <http://arxiv.org/abs/1601.05499>
- [9] Blanchet, J., Chen, X. and Dong, J. (2014). ϵ -Strong simulation for multidimensional stochastic differential equations via rough path analysis. To appear in *Annals of Applied Probability*. <http://arxiv.org/abs/1403.5722>
- [10] Blanchet, J., Dong, J. (2012). Sampling point processes on stable unbounded regions and exact simulation of queues. In *Simulation Conference (WSC), Proceedings of the 2012 Winter* (pp. 1–12). IEEE.
- [11] Blanchet, J., Dong, J. (2015). Perfect sampling for infinite server and loss systems. *Advances in Applied Probability*, 47(3), 761–786.
- [12] Blanchet, J., Dong, J. and Pei, Y. (2015). Perfect sampling of GI/GI/c queues. Submitted to *Queueing Systems: Theory and Applications*.
- [13] Blanchet, J. and Wallwater, A. (2015). Exact sampling of stationary and time reversed queues. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 25(4), Article 26.

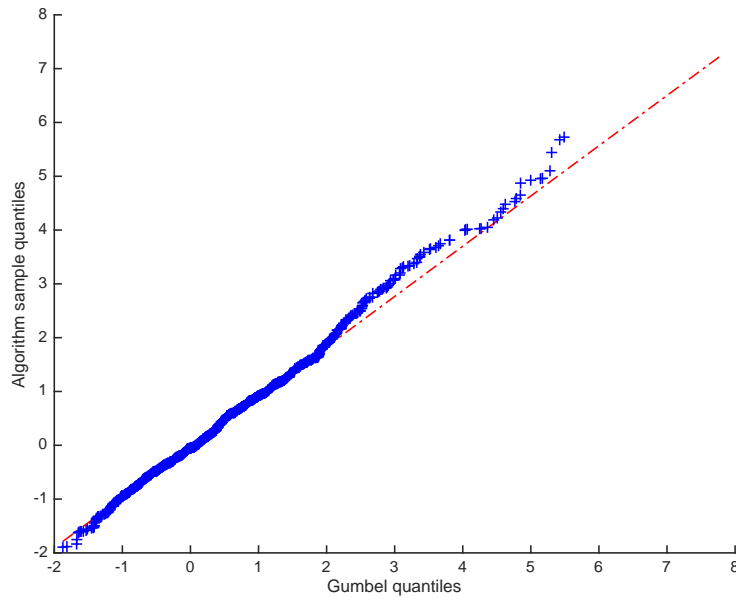


FIGURE 3. The QQ-plot of $M(0.5)$ as generated by our algorithm vs. standard Gumbel

- [14] Blanchet, J. and Sigman, K. (2011). On exact sampling of stochastic perpetuities. *Journal of Applied Probability*, 48A, 165–182.
- [15] Brown, B. M. and Resnick, S. I.. (1977). Extreme values of independent stochastic processes. *Journal of Applied Probability*, 14(4), 732–739.
- [16] Buishand, T. A., de Haan, L., and Zhou, C. (2008). On spatial extremes: With application to a rainfall problem. *The Annals of Applied Statistics*, 2(2), 624–642.
- [17] Dieker, A. B., Mikosch, T. (2015). Exact simulation of Brown-Resnick random fields at a finite number of locations. *Extremes*, 18(2), 301–314.
- [18] Dombry, C., Engelke, S. and Oesting, M. (2015). Exact simulation of max-stable processes. *arXiv:1506.04430*.
- [19] Embrechts, P., Klüppelberg, C., and Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, New York.
- [20] Haan, L. de and Zhou, C. (2008). On extreme value analysis of a spatial process. *RevStat - Statistical Journal*, 6(1), 71–81.

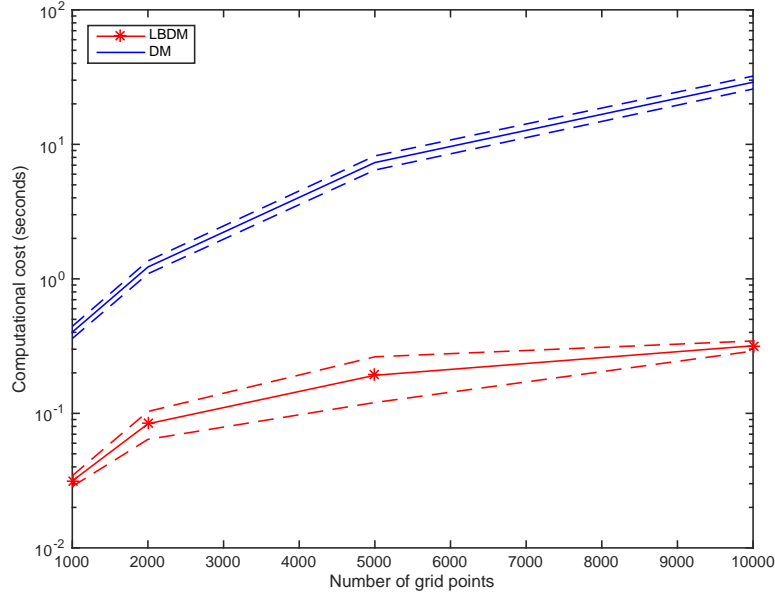


FIGURE 4. Comparison of running time of our algorithm (LBDM) vs. [17] (DM), $H = 1/4$.

- [21] Haan, L. de (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 12(4), 1194–1204.
- [22] Hoffman, Y., Ribak, E. (1991). Constrained realizations of Gaussian fields - a simple algorithm. *The Astrophysical Journal*, vol. 380, pp. L5–L8.
- [23] Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37(5), 2042–2065.
- [24] Kühn, T., Linde, W. (2002). Optimal series representation of fractional Brownian sheets. *Bernoulli*, 8(5), 669–696.
- [25] Oesting, M., Schlather, M., and Zhou, C. (2013). On the normalized spectral representation of max-stable processes on a compact set. *arXiv:1310.1813*.

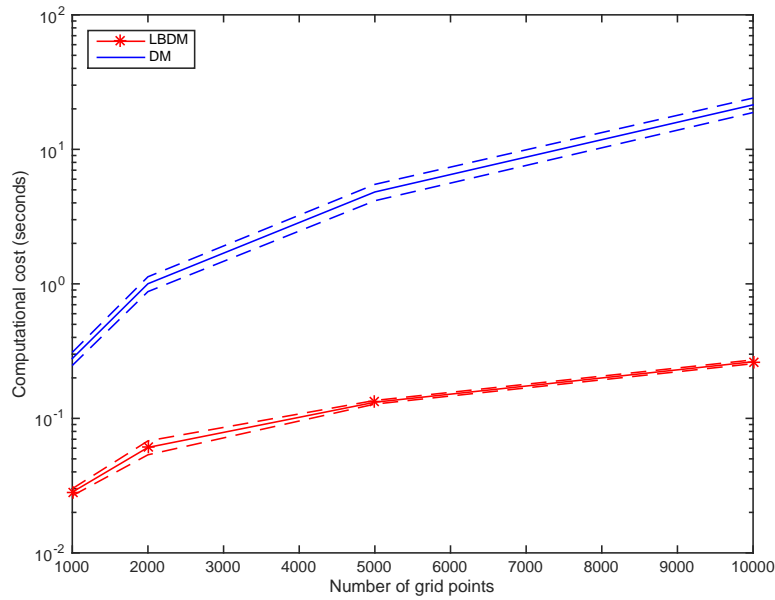


FIGURE 5. Comparison of running time of our algorithm (LBDM) vs. [17] (DM), $H = 1/2$.

- [26] Pollock, M., Johansen, A. M. and Roberts, G. O. (2016). On the exact and ϵ -strong simulation of (jump) diffusions. *Bernoulli*, 22(2), 794–856.
- [27] Schilling, R. L. and Partzsch, L. (2012) *Brownian Motion. An Introduction to Stochastic Processes*. De Gruyter, Berlin/Boston.
- [28] Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5(1), 33–44.
- [29] Smith, R. L. (1990). Max-stable processes and spatial extremes. Unpublished manuscript, Univer.
- [30] Steele, J. M. (2001). *Stochastic Calculus and Financial Applications*. Springer, New York.
- [31] Thibaud, E., Aalto, J., Cooley, D. S., Davison, A. C., and Heikkinen, J. (2015). Bayesian inference for the Brown-Resnick process, with an application to extreme low temperatures. *arXiv:1506.07836*.
- [32] Vaart, A. van der and Wellner, J. A. (1996) *Weak Convergence and Empirical Processes. With Applications to Statistics*. Springer, New York.

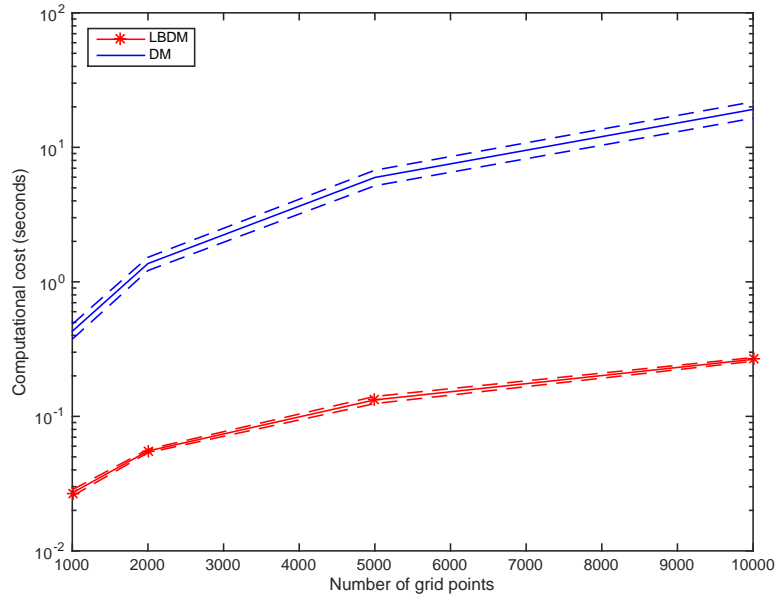


FIGURE 6. Comparison of running time of our algorithm (LBDM) vs. [17] (DM), $H = 3/4$.

INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH DEPARTMENT, SCHOOL OF ENGINEERING AND APPLIED SCIENCES, COLUMBIA UNIVERSITY, 500 WEST 120TH STREET, NEW YORK, NY 10027, U.S.A.

E-mail address: `z12337@columbia.edu`

E-mail address: `jose.blanchet@columbia.edu`

E-mail address: `ton.dieker@ieor.columbia.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COPENHAGEN, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN, DENMARK

E-mail address: `mikosch@math.ku.dk`

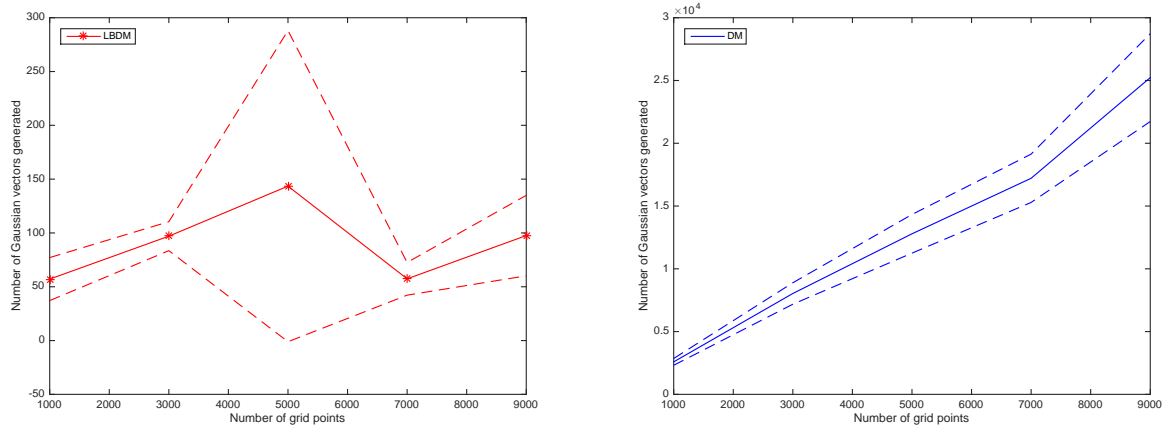


FIGURE 7. Comparison of number of Gaussian vectors generated in our algorithm (LBDM) v.s. [17] (DM), $H = 1/4$.